GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:08:23 ; Search time 872.437 Seconds
(without alignments)
10014.946 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311
Perfect score: 1311
Sequence: 1 ccatggtgtctctatgaaaa.....cgcacatatgtttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1311	100.0	1394	2	AAX07408	Aax07408 Zea mays
	2	1311	100.0	1394	2	AAX07409	Aax07409 Zea mays
	3	1311	100.0	1394	5	AAH76332	Aah76332 Z. mays M
	4	1311	100.0	1394	5	AAH76333	Aah76333 Z. mays M
	5	160.6	12.3	255	5	AAH76340	Aah76340 Z. mays M
	6	146	11.1	158	5	AAH76334	Aah76334 Z. mays M
	7	74	5.6	320	13	ADX48600	Adx48600 Plant ful
	8	74	5.6	624	13	ADX33444	Adx33444 Plant ful
	9	72.8	5.6	2445	13	ADX61650	Adx61650 Plant ful
c	10	71.8	5.5	928	13	ADX34996	Adx34996 Plant ful
	11	71.8	5.5	2537	13	ADX61114	Adx61114 Plant ful
	12	71.8	5.5	3607	13	ADX60218	Adx60218 Plant ful
c	13	70.2	5.4	883	4	AAL15210	Aal15210 Human bre
c	14	70.2	5.4	960	11	ACN85231	Acn85231 Breast ca
	15	69.8	5.3	13400	13	ADR04296	Adr04296 Corn FT h
c	16	69.2	5.3	527	13	ADX09466	Adx09466 Plant ful
c	17	68.4	5.2	1215	13	ADX49916	Adx49916 Plant ful
	18	68	5.2	1092	14	AEB28609	Aeb28609 Maize pol
	19	67.8	5.2	928	13	ADX34996	Adx34996 Plant ful
c	20	67.4	5.1	1260	13	ADX60038	Adx60038 Plant ful
	21	67.2	5.1	2232	13	ADX33787	Adx33787 Plant ful
	22	67.2	5.1	2249	13	ADX45524	Adx45524 Plant ful
	23	67.2	5.1	2657	2	AAZ10551	Aaz10551 DNA seque
c	24	66.4	5.1	6027	2	AAX58751	Aax58751 Maize dul
c	25	66.4	5.1	6027	8	ABX09935	Abx09935 DNA encod
c	26	66.4	5.1	6027	12	ADK12106	Adk12106 cDNA enco
	27	66.2	5.0	439	8	ABX35844	Abx35844 Bovine ES
	28	66	5.0	1376	13	ADO83353	Ado83353 Plant ful
c	29	65.6	5.0	346	4	AAI87279	Aai87279 Human pol
	30	65.6	5.0	2274	13	ADO82110	Ado82110 Plant ful
	31	65.6	5.0	2274	13	ADO82057	Ado82057 Plant ful
c	32	65.6	5.0	8056	8	ABZ10246	Abz10246 Haematopo
c	33	64.6	4.9	9859	13	ADS89714	Ads89714 Oligonucl
c	34	64.4	4.9	9859	13	ADS89440	Ads89440 Oligonucl
c	35	64	4.9	8056	8	ABZ10100	Abz10100 Haematopo
	36	63.8	4.9	431	8	ABX44556	Abx44556 Bovine ES
c	37	63.6	4.9	335913	5	AAI61371	Aai61371 Soybean 2
c	38	63.6	4.9	335913	5	AAI61372	Aai61372 Soybean 2
	39	63.4	4.8	12968	4	AAS45494	Aas45494 Chemicall
	40	63.4	4.8	12968	4	AAS46779	Aas46779 Tumour su
	41	63.4	4.8	12968	6	ABL34085	Abl34085 Human imm
	42	63.4	4.8	12968	6	ABK28424	Abk28424 DNA trans
	43	63.4	4.8	12968	6	ABN80285	Abn80285 Human che
	44	63.2	4.8	1340	14	ADZ71035	Adz71035 Human chr
	45	63	4.8	7624	6	ABL34113	Abl34113 Human imm

ALIGNMENTS

RESULT 1
AAX07408
ID AAX07408 standard; DNA; 1394 BP.
XX
AC AAX07408;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW Ms45; male; tissue-preferred; regulatory region; plant cells;
KW plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
PN WO9859061-A1.
XX
PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98WO-US012895.
XX
PR 23-JUN-1997; 97US-00880499.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
DR WPI; 1999-105628/09.
XX
PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
PS Claim 2; Page 22-23; 39pp; English.
XX
CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue- preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile

```
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 1311;  DB 2;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-269;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

```
RESULT 2
AAX07409
ID   AAX07409 standard; DNA; 1394 BP.
XX
AC   AAX07409;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW   Ms45; male; tissue-preferred; regulatory region; plant cells;
KW   plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS   Zea mays.
XX
PN   WO9859061-A1.
XX
PD   30-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US012895.
XX
PR   23-JUN-1997;   97US-00880499.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman GA,  Kendall TL;
XX
DR   WPI; 1999-105628/09.
XX
PT   New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT   - useful in mediating plant fertility, especially hybrid seed production.
XX
PS   Claim 3; Page 23-24; 39pp; English.
XX
CC   The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC   region. It may be used in the construction of a vector for a method of
CC   producing exogenous genes in a male tissue- preferred manner, which is
CC   useful in restoring or conferring fertility, such as in hybrid seed
CC   production. In conferring fertility, a monocot/dicot plant is transformed
CC   with the exogenous nucleotide sequence (a male sterility gene, preferably
CC   Ms45), which encodes a product selected from auxins, rolB and diptheria
CC   toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 1311;  DB 2;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-269;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
Db       1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy      61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
Db      61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy     121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
Db     121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy     181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
Db     181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy     241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
Db     241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy     301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
Db     301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy     361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
Db     361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy     421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
Db     421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy     481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
Db     481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy     541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
Db     541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy     601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
Db     601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy     661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
Db     661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy     721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
Db     721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy     781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
Db     781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy     841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
Db     841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy     901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
Db     901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy     961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
Db     961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy    1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
Db    1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy    1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
Db    1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy    1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
Db    1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy    1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
Db    1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy    1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
Db    1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 3
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 1311; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.2e-269;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
Db       1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy      61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
```

```
Db        61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy       121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Db       121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy       181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Db       181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy       241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Db       241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy       301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Db       301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy       361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Db       361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy       421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Db       421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy       481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Db       481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy       541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Db       541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy       601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Db       601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy       661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Db       661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy       721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Db       721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy       781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Db       781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy       841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Db       841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy       901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Db       901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy       961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Db       961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy      1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Db      1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy      1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Db      1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy      1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Db      1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy      1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Db      1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy      1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311

Db      1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 4
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
XX
AC AAH76333;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 100.0%; Score 1311; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.2e-269;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
```

```
Qy       181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
Db       181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy       241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
Db       241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy       301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
Db       301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy       361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
Db       361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy       421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
Db       421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy       481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
Db       481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy       541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
Db       541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy       601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
Db       601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy       661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
Db       661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy       721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
Db       721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy       781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
Db       781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy       841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
Db       841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy       901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
Db       901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy       961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
Db       961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy      1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
Db      1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy      1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
Db      1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy      1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
Db      1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy      1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
Db      1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy      1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
Db      1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 5
AAH76340
ID AAH76340 standard; DNA; 255 BP.
XX
AC AAH76340;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 promoter fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; promoter; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Example 5; Fig 8; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a Z. mays Ms45 promoter fragment
XX
SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 12.3%; Score 160.6; DB 5; Length 255;
Best Local Similarity 97.6%; Pred. No. 1.5e-24;
Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1145 CGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGT 1204
Db         5 CCGCGGATCCCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGT 64

Qy      1205 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 1264
Db        65 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 124

Qy      1265 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
Db       125 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171
```

RESULT 6
AAH76334
ID AAH76334 standard; DNA; 158 BP.
XX
AC AAH76334;

```
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 5; Page 47; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a DNA fragment -38 to -195 bases upstream of
CC   the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC   nucleotide sequence
XX
SQ   Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

  Query Match             11.1%;  Score 146;  DB 5;  Length 158;
  Best Local Similarity   99.4%;  Pred. No. 1.7e-21;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 1213
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy       1214 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 1273
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120

Qy       1274 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
             |||||||||||||||||||||||||||||||||||||
Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158


RESULT 7
ADX48600
ID   ADX48600 standard; cDNA; 320 BP.
XX
AC   ADX48600;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 23340.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 23340; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This sequence represents a plant full length insert
CC   polynucleotide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 320 BP; 89 A; 53 C; 66 G; 112 T; 0 U; 0 Other;

  Query Match              5.6%;  Score 74;  DB 13;  Length 320;
  Best Local Similarity   73.1%;  Pred. No. 4.6e-06;
  Matches   95;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy        791 TTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAA 850
              |||| || |||||||||| || |   ||| ||||||   ||||| |||||   |||  ||
Db         62 TTTTTAGGGCTAGTTTGGAAATCTCATTTTTTTCAAGAGATTTTTATTTTCCTAAAGGAA 121

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
              |||||||||||||   || | |||||||||   | ||| | ||| ||||||| | | || |
Db        122 ATTAGTTTATTTTTCTTTGAGAAAATAGAATTCCCTTGGGAAATTAGAGTTTCTAAACGA 181

Qy        911 GCCCTAGAAT 920
              |||||  | |
Db        182 GCCCTTAATT 191
```

```
RESULT 8
ADX33444
ID   ADX33444 standard; cDNA; 624 BP.
XX
AC   ADX33444;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 16264.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 16264; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This sequence represents a plant full length insert
CC   polynucleotide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 624 BP; 170 A; 119 C; 146 G; 189 T; 0 U; 0 Other;

  Query Match             5.6%;  Score 74;  DB 13;  Length 624;
  Best Local Similarity   73.1%;  Pred. No. 5.4e-06;
  Matches   95;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy        791 TTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAA 850
              |||| || |||||||||| || |   ||| ||||||   ||||| |||||   ||| ||
Db        374 TTTTTAGGGCTAGTTTGGAAATCTCATTTTTTTTCAAGAGATTTTTATTTTCCTAAAGGAA 433

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
              |||||||||||||    || | |||||||||  | ||| | ||| ||||||| | | || |
Db        434 ATTAGTTTATTTTTCTTTGAGAAAATAGAATTCCCTTGGGAAATTAGAGTTTCTAAACGA 493

Qy        911 GCCCTAGAAT 920
              |||||  | |
Db        494 GCCCTTAATT 503


RESULT 9
ADX61650
ID   ADX61650 standard; cDNA; 2445 BP.
XX
AC   ADX61650;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 32493.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 32493; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
```

CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 2445 BP; 666 A; 470 C; 513 G; 796 T; 0 U; 0 Other;

Query Match 5.6%; Score 72.8; DB 13; Length 2445;
Best Local Similarity 73.6%; Pred. No. 1.4e-05;
Matches 106; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

```
Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTT-TCAAAGAATTTTGATTTTTTCAAAAAAA 850
              | | || |||||||||  |||||| ||| ||  ||||  |||||  ||||||  || ||||
Db       1623 TCTTAGGGCTAGTTTGAGAACCCTTTTTTTTCCCAAAAGATTTTCATTTTTCAAAGAAAA 1682

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
              |||||||  |||||  |||  | |||||||  || |  ||||  ||||||||  |||  ||  ||||
Db       1683 ATTAGTTCATTTTTCCTTGAGAAAATAGGAATCCCTTAAAAAAATAGTGTTTTCAAACTA 1742

Qy        911 GCCCTAGAATGTTTTCCCAATAAA 934
              |||||  |   ||||    || ||
Db       1743 GCCCTTAATGTTTTTTTTCATGAA 1766
```

RESULT 10
ADX34996/c
ID ADX34996 standard; cDNA; 928 BP.
XX
AC ADX34996;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polynucleotide seqid 17816.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 17816; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 928 BP; 245 A; 209 C; 219 G; 255 T; 0 U; 0 Other;

Query Match 5.5%; Score 71.8; DB 13; Length 928;
Best Local Similarity 78.0%; Pred. No. 1.8e-05;
Matches 99; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

```
Qy        790 TTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
              |||||||||||| |||||||||  |||  || ||  || | ||||  |||||   || |||
Db        804 TTTTTAAGAGCTGGTTTGGCAA-CCTCATTTTTCTAAGGGATTTCTATTTTACTAAGAAA 746

Qy        850 AATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACT 909
              |||||||| ||||||||||   |||||| |||   ||| ||||||||||||| ||| | |
Db        745 AATTAGTTAATTTTCTCTTAGAAAAATAAAAATGTCTTGGAAAAATAGAGTTACCAAATT 686

Qy        910 AGCCCTA 916
              | |||||
Db        685 ACCCCTA 679
```

RESULT 11
ADX61114
ID ADX61114 standard; cDNA; 2537 BP.
XX
AC ADX61114;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polynucleotide seqid 31957.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.

PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 31957; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 2537 BP; 685 A; 484 C; 524 G; 844 T; 0 U; 0 Other;

Query Match 5.5%; Score 71.8; DB 13; Length 2537;
Best Local Similarity 73.4%; Pred. No. 2.2e-05;
Matches 105; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

```
Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              | | || |||||||||  |||||| ||| |  ||||  ||||| ||||||  || |||||
Db       1799 TCTTAGGGCTAGTTTGAGAACCCT-TTTTTCCCAAAAGATTTTCATTTTTCAAAGAAAAA 1857

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              ||||||  |||||  ||| | |||||||  || | ||||  |||||||| |||  || |||||
Db       1858 TTAGTTCATTTTTCCTTGAGAAAATAGGAATCCCTTAAAAAAATAGTGTTTTCAAACTAG 1917

Qy        912 CCCTAGAATGTTTTCCCAATAAA 934
              ||||  |  ||||   || ||
Db       1918 CCCTTAATGTTTTTTTTCATGAA 1940
```

RESULT 12
ADX60218
ID ADX60218 standard; cDNA; 3607 BP.
XX
AC ADX60218;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polynucleotide seqid 31061.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 31061; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 3607 BP; 915 A; 695 C; 718 G; 1279 T; 0 U; 0 Other;

Query Match 5.5%; Score 71.8; DB 13; Length 3607;
Best Local Similarity 73.4%; Pred. No. 2.4e-05;
Matches 105; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

```
Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              | | || |||||||||  |||||| ||| |  ||||  ||||| ||||||  || |||||
Db       2871 TCTTAGGGCTAGTTTGAGAACCCT-TTTTTCCCAAAAGATTTTCATTTTTCAAAGAAAAA 2929

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              ||||||  |||||  ||| | |||||||  || | ||||  |||||||| |||  || |||||
Db       2930 TTAGTTCATTTTTCCTTGAGAAAATAGGAATCCCTTAAAAAAATAGTGTTTTCAAACTAG 2989

Qy        912 CCCTAGAATGTTTTCCCAATAAA 934
              ||||  |  ||||   || ||
Db       2990 CCCTTAATGTTTTTTTTCATGAA 3012
```

RESULT 13
AAL15210/c
ID AAL15210 standard; cDNA; 883 BP.
XX
AC AAL15210;
XX
DT 07-DEC-2001 (first entry)
XX

DE Human breast cancer expressed polynucleotide 7667.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS Claim 1; Page 1378; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 883 BP; 322 A; 18 C; 23 G; 396 T; 0 U; 124 Other;

Query Match 5.4%; Score 70.2; DB 4; Length 883;
Best Local Similarity 39.2%; Pred. No. 3.8e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

```
Qy        419 TGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCA 478
              |||  |       |   |||   ||  |  |   |     |      |||   | | |
Db        873 TGTTNANNNANACCTANACCTTTAAANATNTNNNANTTNAANNAAAAAAATATTTAAANN 814

Qy        479 CTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAA 538
                ||   | | |   | |  ||  |   | |   | |      |  |        ||||
Db        813 TANAAATTTTTTNAAAAACTAATTNTAANNATANTNAATTTTTNTAAAAAAANNNAAAAA 754

Qy        539 AACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACA 598
                ||   | | |          | |     |  ||     |||| ||| || |   |
Db        753 TTAAANNTTTTNTAANTTATTAAACCAAAATTTTTTTTTTAAAAAAATTTTTTAAAANTT 694

Qy        599 AATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTG 658
              |||  |  |  |     |       |    | | |||    |||  |  |   |  ||||
Db        693 AATAANTTNAATTTTTTTTAATNAAAAAAAAAAATTTNAAATTTTAANCAANTNTTTTTT 634

Qy        659 TTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATA 718
              ||  | |   || | | | ||| |      | ||     |  |||| |    || ||  |
Db        633 TTTNTTTNNATAAAAAAANTTTTTAAATAAAAAANNANTTTTTTTTATNAANATATTTAA 574

Qy        719 AAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAA 778
              |||  |    |   | ||| ||  | | |  |     |    |  || |  | |||  |
Db        573 AAAAAAANNTTTTTTTTTTTATANTTTATAAAAAATTTTATTTNTNTTTTTCTNAAAANA 514

Qy        779 TCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATT 838
                |||||  ||    |  |    |       |          ||    |||   ||
Db        513 AAAAAAAATTTTTTNAAANTNTTTTTAAAACCTTNAATTNAAANANAATTTTTTTNNNTT 454
```

```
Qy        839 TTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGA 898
              ||||||||  ||| ||||   | ||||||  || || |      |  |||||  |
Db        453 NNANATTAAAAAATTTATTTTTTTTANTTCTATAAATTAAAANAANANNAATAAAATAAT 394

Qy        899 GTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATT 958
               ||   | | ||    || | |  |       |  || ||| ||          | |||| |
Db        393 TTTNAAAAAATAATTTTATANTTATNNAANANAAATTTAAAANTNAANANAAAAATTTTA 334

Qy        959 TGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCT 1018
                  |      |||| | || || |   | |   |      ||   |||    | |
Db        333 NTTTAAATAAATAAAATAANTTTAAACNTCATTTNTAATTTTATAATTAAAAAAAAAATT 274

Qy       1019 ATTT 1022
                |||
Db        273 NTTT 270
```

RESULT 14
ACN85231/c
ID ACN85231 standard; DNA; 960 BP.
XX
AC ACN85231;
XX
DT 02-DEC-2004 (first entry)
XX
DE Breast cancer related marker, seq id 6381.
XX
KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
OS Homo sapiens.
XX
PN US2003099974-A1.
XX
PD 29-MAY-2003.
XX
PF 18-JUL-2002; 2002US-00198846.
XX
PR 18-JUL-2001; 2001US-0306220P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2003-787014/74.
XX
PT Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
PS Disclosure; SEQ ID NO 6381; 36pp; English.
XX
CC The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
SQ Sequence 960 BP; 340 A; 33 C; 39 G; 421 T; 0 U; 127 Other;

Query Match 5.4%; Score 70.2; DB 11; Length 960;
Best Local Similarity 39.2%; Pred. No. 3.9e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

```
Qy        419 TGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCA 478
              |||  |       |   |||   ||  |  |   |     |      |||   | | |
```

```
Db        933 TGTTNANNNANACCTANACCTTTAAANATNTNNNANTTNAANNAAAAAAATATTTAAANN 874

Qy        479 CTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAA 538

Db        873 TANAAATTTTTTNAAAAACTAATTNTAANNATANTNAATTTTTNTAAAAAAANNNAAAAA 814

Qy        539 AACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACA 598

Db        813 TTAAANNTTTTNTAANTTATTAAACCAAAATTTTTTTTTTAAAAAAATTTTTTAAAANTT 754

Qy        599 AATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTG 658

Db        753 AATAANTTNAATTTTTTTTAATNAAAAAAAAAAATTTNAAATTTTAANCAANTNTTTTTT 694

Qy        659 TTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATA 718

Db        693 TTTNTTTNNATAAAAAAANTTTTTAAATAAAAAANNANTTTTTTTTATNAANATATTTAA 634

Qy        719 AAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAA 778

Db        633 AAAAAAANNTTTTTTTTTTTATANTTTATAAAAAATTTTATTTNTNTTTTTCTNAAAANA 574

Qy        779 TCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATT 838

Db        573 AAAAAAAATTTTTTNAAANTNTTTTTAAAACCTTNAATTNAAANANAATTTTTTNNNTT 514

Qy        839 TTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGA 898

Db        513 NNANATTAAAAAATTTATTTTTTTTANTTCTATAAATTAAAANAANANNAATAAAATAAT 454

Qy        899 GTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATT 958

Db        453 TTTNAAAAAATAATTTTATANTTATNNAANANAAATTTAAAANTNAANANAAAAATTTTA 394

Qy        959 TGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCT 1018

Db        393 NTTTAAATAAATAAAATAANTTTAAACNTCATTTNTAATTTTATAATTAAAAAAAAAATT 334

Qy       1019 ATTT 1022

Db        333 NTTT 330
```

RESULT 15
ADR04296
ID ADR04296 standard; DNA; 13400 BP.
XX
AC ADR04296;
XX
DT 04-NOV-2004 (first entry)
XX
DE Corn FT homologue nucleotide sequence SEQ ID NO:63.
XX
KW flowering locus T gene; FT; terminal flower; TFL; Apetala3; Ap3; plant;
KW floral development; plant sterility; plant fertility; flowering time;
KW plant growth rate; inflorescence architecture; tissue culture morphology;
KW cell division; FT homologue; gene; ds.
XX
OS Zea mays.
XX
PN WO2004067723-A2.
XX
PD 12-AUG-2004.
XX
PF 29-JAN-2004; 2004WO-US002422.
XX
PR 30-JAN-2003; 2003US-00343477.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Danilevskaya O, Hermon P, Bruggemann E, Shirbroun D, Ananiev E;
PI Rafalski JA, Sakai H, Cahoon E, Cahoon R, Klein T;
XX
DR WPI; 2004-580996/56.
XX
PT New polynucleotides, specifically nucleic acid fragments encoding
PT flowering locus T gene (FT) or terminal flower (TFL), or Apetala3 (Ap3)
PT homologs, useful for floral development, e.g. engineering plant flowering
PT time.
XX
PS Claim 6; SEQ ID NO 63; 109pp; English.
XX
CC The present invention describes an isolated polynucleotide comprising a
CC first, second, third, fourth or fifth nucleotide sequence, or their
CC complement encoding a polypeptide either having flowering locus T gene
CC (FT), terminal flower (TFL), or Apetala3 (Ap3) homologue activity. Also
CC described: (1) a vector comprising the polynucleotide; (2) a recombinant
CC DNA construct comprising the polynucleotide; (3) transforming a cell by
CC transforming a cell with the polynucleotide; (4) a cell comprising the
CC recombinant DNA construct; (5) producing a plant comprising transforming
CC a plant cell with the polynucleotide, and regenerating a plant from the
CC transformed plant cell; (6) a plant comprising the recombinant DNA
CC construct; (7) a seed comprising the recombinant DNA construct; (8) an
CC isolated polynucleotide comprising a first nucleotide sequence, where the
CC first nucleotide sequence contains at least 30 nucleotides, and where the
CC first nucleotide sequence is comprised by another polynucleotide, where
CC the other polynucleotide includes the second, third, fourth, fifth or
CC sixth nucleotide sequence; (9) an isolated polypeptide having FT or Ap3
CC homologue activity, as described above; and (10) isolating a polypeptide
CC encoded by the polynucleotide comprising isolating the polypeptide from a
CC cell containing a recombinant DNA construct comprising the polynucleotide
CC operably linked to a regulatory sequence. The polynucleotides are useful
CC for floral development, e.g. engineering plant sterility/fertility,
CC flowering time, plant growth rate, inflorescence architecture, and tissue
CC culture morphology and the rate of cell division to enhance
CC transformation. The present sequence represents an FT homologue
CC nucleotide sequence from the present invention.
XX
SQ Sequence 13400 BP; 3962 A; 2948 C; 2695 G; 3795 T; 0 U; 0 Other;

  Query Match   5.3%; Score 69.8; DB 13; Length 13400;
  Best Local Similarity 73.6%; Pred. No. 9e-05;
  Matches   89; Conservative   0; Mismatches   32; Indels   0; Gaps   0;

```
Qy        811 ACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTA 870

Db       8559 ACCATTTTATTTTCAAGAGGTTTTTATTTTATCAAGAAAAATTAGTTCATTTTCTCTTGG 8618

Qy        871 TAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAA 930

Db       8619 AAAAATAAAAATCCATTAGAAAAATGGGGTTGTCAAACTAGTCCTTATTTAGTTTTCCAT 8678

Qy        931 T 931

Db       8679 T 8679
```

Search completed: March 5, 2006, 18:11:32
Job time : 875.437 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:38:03 ; Search time 5721.35 Seconds
(without alignments)
13025.213 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311
Perfect score: 1311
Sequence: 1 ccatggtgtctctatgaaaa.....cgcacatatgtttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1311	100.0	1394	6	BD062176	BD062176 Male tiss
2	1311	100.0	1394	6	BD062177	BD062177 Male tiss
3	1311	100.0	1394	6	AX224394	AX224394 Sequence
4	1311	100.0	1394	6	AX224395	AX224395 Sequence
5	1311	100.0	3343	15	AF360356	AF360356 Zea mays
6	160.6	12.3	255	6	AX224402	AX224402 Sequence
7	146	11.1	158	6	AX224396	AX224396 Sequence
8	77.2	5.9	188283	14	AC155584	AC155584 Zea mays
9	77	5.9	186199	14	AC147602	AC147602 Zea mays
10	74.4	5.7	161616	14	CR936840	CR936840 Danio rer
11	73.4	5.6	25769	2	AC117267	AC117267 Dictyoste
c 12	72.4	5.5	172293	14	AC155379	AC155379 Zea mays
c 13	71.6	5.5	201985	14	AC155618	AC155618 Zea mays
c 14	71.2	5.4	713	2	AF034389	AF034389 Plasmodiu
c 15	71.2	5.4	3576	2	DDU86962	U86962 Dictyosteli
16	71	5.4	317137	15	AY664413	AY664413 Zea mays
c 17	70.8	5.4	149598	5	BX957346	BX957346 Zebrafish
c 18	70.8	5.4	156452	14	AC155376	AC155376 Zea mays
c 19	70.4	5.4	347582	2	PFMAL4P1	AL034557 Plasmodiu
c 20	70.2	5.4	883	6	CQ422641	CQ422641 Sequence
c 21	70	5.3	241	10	BV119878	BV119878 PZA01377
c 22	70	5.3	247	10	BV119874	BV119874 PZA01377
c 23	70	5.3	248	10	BV119877	BV119877 PZA01377
24	69.8	5.3	13400	6	CQ855052	CQ855052 Sequence
c 25	69.6	5.3	134692	14	AC151050	AC151050 Zea mays
c 26	69.4	5.3	110000	15	CR382132_17	Continuation (18 o
27	69.2	5.3	112468	14	AC149836	AC149836 Zea mays
28	69.2	5.3	266544	2	AC116956	AC116956 Dictyoste
c 29	69	5.3	148120	14	AC155517	AC155517 Zea mays
c 30	68.8	5.2	7347	1	AF211124	AF211124 Carsonell
c 31	68.6	5.2	186199	14	AC147602	AC147602 Zea mays
c 32	68.4	5.2	213	10	BV119875	BV119875 PZA01377
c 33	68.4	5.2	216	10	BV119871	BV119871 PZA01377
c 34	68.4	5.2	224	10	BV119882	BV119882 PZA01377
c 35	68.4	5.2	245	10	BV119881	BV119881 PZA01377
c 36	68.4	5.2	246	10	BV119869	BV119869 PZA01377
37	68.2	5.2	346296	15	AF090447	AF090447 Zea mays
38	68	5.2	1092	6	CS123057	CS123057 Sequence
39	68	5.2	110000	14	PFMAL7P1_02	Continuation (3 of
40	67.8	5.2	227495	14	AC148480	AC148480 Zea mays
c 41	67.6	5.2	351	10	BV110530	BV110530 PZA01059
42	67.6	5.2	110000	15	AY664419_2	Continuation (3 of
43	67.6	5.2	331039	2	AC116979	AC116979 Dictyoste
c 44	67.4	5.1	170	10	BV119873	BV119873 PZA01377
c 45	67.2	5.1	155263	14	BX927284	BX927284 Danio rer

ALIGNMENTS

RESULT 1
BD062176
LOCUS BD062176 1394 bp DNA linear PAT 27-AUG-2002
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062176
VERSION BD062176.1 GI:22607781
KEYWORDS JP 2001520523-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC ,C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES Location/Qualifiers
source 1. .1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 7.2e-228;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 2
BD062177
LOCUS BD062177 1394 bp DNA linear PAT 27-AUG-2002
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062177
VERSION BD062177.1 GI:22607782
KEYWORDS JP 2001520523-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/2
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC ,C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES Location/Qualifiers
source 1. .1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 7.2e-228;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
```

```
Qy      241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy      301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy      361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy      421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy      481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy      541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy      601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy      661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy      721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy      781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy      841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy      901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy      961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy     1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy     1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy     1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy     1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy     1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

```
RESULT 3
AX224394
LOCUS       AX224394               1394 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 1 from Patent WO0160997.
ACCESSION   AX224394
VERSION     AX224394.1  GI:15554636
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 1 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1394
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match           100.0%;  Score 1311;  DB 6;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-228;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
```

```
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 4
AX224395
LOCUS       AX224395               1394 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 2 from Patent WO0160997.
ACCESSION   AX224395
VERSION     AX224395.1  GI:15554637
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 2 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1394
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             100.0%;  Score 1311;  DB 6;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-228;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 5
AF360356
LOCUS       AF360356               3343 bp    DNA     linear   PLN 12-MAY-2001
DEFINITION  Zea mays male fertility protein (Ms45) gene, complete cds.
ACCESSION   AF360356
VERSION     AF360356.1  GI:14028756
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3343)
  AUTHORS   Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
  TITLE     Cloning of Ms45, a gene required for male ferility from Zea mays
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3343)
  AUTHORS   Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
            Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
            IA 50131-1004, USA
FEATURES             Location/Qualifiers
     source          1. .3343
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /chromosome="9L"
     gene            <1392. .>2942
                     /gene="Ms45"
     mRNA            join(<1392. .1768,1898. .2182,2280. .2447,2534. .>2942)
                     /gene="Ms45"
                     /product="male fertility protein"
     CDS             join(1392. .1768,1898. .2182,2280. .2447,2534. .2942)
                     /gene="Ms45"
                     /codon_start=1
                     /product="male fertility protein"
                     /protein_id="AAK52489.1"
                     /db_xref="GI:14028757"
                     /translation="MEKRNLQWRRGRDGIVQYPHLFFAALALALLVADPFGLSPLAEV
                     DYRPVKHELAPYGEVMGSWPRDNASRLRRGRLEFVGEVFGPESIEFDLQGRGPYAGLA
                     DGRVVRWMGEEAGWETFAVMNPDWSEEVCANGVNSTTRKQHEKEEFCGRPLGLRFHGE
                     TGELYVADAYYGLMVVGQSGGVASSVAREADGDPIRFANDLDVHRNGSVFFTDTSMRY
                     SRKDHLNILLEGEGTGRLLRYDPETSGVHVVLKGLVFPNGVQISEDHQFLLFSETTNC
                     RIMRYWLEGPRAGEVEVFANLPGFPDNVRSNGRGQFWVAIDCCRTPAQEVFAKRPWLR
                     TLYFKFPLSLKVLTWKAARRMHTVLALLDGEGRVVEVLEDRGHEVMKLVSEVREVGRK
                     LWIGTVAHNHIATIPYPLED"
ORIGIN

  Query Match           100.0%;  Score 1311;  DB 15;  Length 3343;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-228;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 6
AX224402
LOCUS       AX224402                 255 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 9 from Patent WO0160997.
ACCESSION   AX224402
VERSION     AX224402.1  GI:15554644
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 9 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .255
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             12.3%;  Score 160.6;  DB 6;  Length 255;
  Best Local Similarity   97.6%;  Pred. No. 4.5e-19;
  Matches  163;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1145 CGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGT 1204
              | || | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 CCGCGGATCCCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGT 64

Qy       1205 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 124

Qy       1265 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        125 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171


RESULT 7
AX224396
LOCUS       AX224396                 158 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 3 from Patent WO0160997.
ACCESSION   AX224396
VERSION     AX224396.1  GI:15554638
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 3 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .158
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             11.1%;  Score 146;  DB 6;  Length 158;
  Best Local Similarity   99.4%;  Pred. No. 2.2e-16;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 1213
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy       1214 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 1273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120

Qy       1274 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
              |||||||||||||||||||||||||||||||||||||
Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158


RESULT 8
AC155584
LOCUS       AC155584               188283 bp    DNA     linear   HTG 25-JAN-2005
DEFINITION  Zea mays strain B73 clone ZMMBBc0196I14, *** SEQUENCING IN PROGRESS
            ***, 20 unordered pieces.
ACCESSION   AC155584
VERSION     AC155584.2  GI:58082443
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 188283)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Consortium for Maize Genomics - BAC skim sequencing and assembly
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 188283)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
REFERENCE   3  (bases 1 to 188283)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
COMMENT     On Jan 25, 2005 this sequence version replaced gi:57863105.
            -------------- Trace submission
                Center_name: TIGR
                Seq_lib_id: ZGFX
            -------------- Project informaton
                Web site: http://www.tigr.org/tdb/tgi/maize/
                Contact: maize@tigr.org
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 20 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
```

```
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    22133: contig of 22133 bp in length
            *    22134    22233: gap of unknown length
            *    22234    27964: contig of 5731 bp in length
            *    27965    28064: gap of unknown length
            *    28065    35160: contig of 7096 bp in length
            *    35161    35260: gap of unknown length
            *    35261    49754: contig of 14494 bp in length
            *    49755    49854: gap of unknown length
            *    49855    96501: contig of 46647 bp in length
            *    96502    96601: gap of unknown length
            *    96602    98675: contig of 2074 bp in length
            *    98676    98775: gap of unknown length
            *    98776   103999: contig of 5224 bp in length
            *   104000   104099: gap of unknown length
            *   104100   123477: contig of 19378 bp in length
            *   123478   123577: gap of unknown length
            *   123578   126464: contig of 2887 bp in length
            *   126465   126564: gap of unknown length
            *   126565   134044: contig of 7480 bp in length
            *   134045   134144: gap of unknown length
            *   134145   142804: contig of 8660 bp in length
            *   142805   142904: gap of unknown length
            *   142905   154520: contig of 11616 bp in length
            *   154521   154620: gap of unknown length
            *   154621   159716: contig of 5096 bp in length
            *   159717   159816: gap of unknown length
            *   159817   160976: contig of 1160 bp in length
            *   160977   161076: gap of unknown length
            *   161077   162085: contig of 1009 bp in length
            *   162086   162185: gap of unknown length
            *   162186   167405: contig of 5220 bp in length
            *   167406   167505: gap of unknown length
            *   167506   171505: contig of 4000 bp in length
            *   171506   171605: gap of unknown length
            *   171606   173244: contig of 1639 bp in length
            *   173245   173344: gap of unknown length
            *   173345   184423: contig of 11079 bp in length
            *   184424   184523: gap of unknown length
            *   184524   188283: contig of 3760 bp in length.
FEATURES             Location/Qualifiers
     source          1. .188283
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBc0196I14"
     gap             22134. .22233
                     /estimated_length=unknown
     gap             27965. .28064
                     /estimated_length=unknown
     gap             35161. .35260
                     /estimated_length=unknown
     gap             49755. .49854
                     /estimated_length=unknown
     gap             96502. .96601
                     /estimated_length=unknown
     gap             98676. .98775
                     /estimated_length=unknown
     gap             104000. .104099
                     /estimated_length=unknown
     gap             123478. .123577
                     /estimated_length=unknown
     gap             126465. .126564
                     /estimated_length=unknown
     gap             134045. .134144
                     /estimated_length=unknown
     gap             142805. .142904
                     /estimated_length=unknown
     gap             154521. .154620
                     /estimated_length=unknown
     gap             159717. .159816
                     /estimated_length=unknown
     gap             160977. .161076
                     /estimated_length=unknown
     gap             162086. .162185
                     /estimated_length=unknown
     gap             167406. .167505
                     /estimated_length=unknown
     gap             171506. .171605
                     /estimated_length=unknown
     gap             173245. .173344
                     /estimated_length=unknown
     gap             184424. .184523
                     /estimated_length=unknown
ORIGIN

  Query Match             5.9%;  Score 77.2;  DB 14;  Length 188283;
  Best Local Similarity   80.5%;  Pred. No. 0.00021;
  Matches  103;  Conservative    0;  Mismatches   23;  Indels    2;  Gaps    1;

Qy        789 TTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAA 848
              || |||||| |||||||| |||||||| ||| || | ||||| ||||| | || |
Db      35512 TTGTTTAAGGGCTAGTTTAAGGACCCTGTT--TTTAAAGGGATTTTTATTTTCTTAAGAG 35569

Qy        849 AAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGAC 908
              |||||||||||||| || | |||||||||| | |||||||||||| |||| ||| ||
Db      35570 AAATTAGTTTATTTTTCATAGTAAAATAGAAATCTCTTAGAAAAATATAGTTCCCAAAC 35629

Qy        909 TAGCCCTA 916
              ||||||||
Db      35630 TAGCCCTA 35637


RESULT 9
AC147602
LOCUS       AC147602               186199 bp    DNA     linear   HTG 17-AUG-2004
DEFINITION  Zea mays clone ZMMBBc0334A01, *** SEQUENCING IN PROGRESS ***, 6
            ordered pieces.
ACCESSION   AC147602
VERSION     AC147602.5  GI:51315585
KEYWORDS    HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 186199)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Butler,E., Wing,R., Bharti,A.K.
            and Messing,J.
  TITLE     Zea mays, clone ZMMBBc0334A01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 186199)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (19-DEC-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 186199)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Butler,E., Wing,R., Bharti,A.K.,
            Messing,J., Abouelleil,A., Allen,N., Anderson,M., Anderson,S.,
            Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S.,
            Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P.,
            FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L.,
            Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P.,
            Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Aug 17, 2004 this sequence version replaced gi:49658659.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@broad.mit.edu
            --------------
                Bharti, AK and Messing, J: The Plant Genome Initiative at
            Rutgers, Waksman Institute, Rutgers, The State University of New
            Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            (http://pgir.rutgers.edu)
                Butler, E and Wing, R: Arizona Genomics Institute, Biological
            Sciences West, 448A, P.O. Box 210088, University of Arizona,
            Tucson, AZ 85721, USA (http://www.genome.arizona.edu)
            -------------- Project Information
                Center project name: L30003
                Center clone name: 334_A_1
            -------------- Consensus Information
            This consensus is derived from a shotgun assembly that has been
            manually curated. It is the best representation of the BAC that we
            can generate without further laboratory work. The draft assembly
            has been edited, and if possible, ends identified by vector as well
            as by BAC end sequences, and contigs ordered and oriented. Bases
            that are not Ns are either above Q20 or manually edited. This
            assembly was performed with Arachne (Genome Res. 2002 12: 177-189;
            Genome Res. 2003 13: 91-96). All trace files for this project are
            available at the NCBI trace repository
            (http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?). An exact list of
            reads used in this assembly are available at
            http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   100617: contig of 100617 bp in length
            *  100618   100717: gap of unknown length
            *  100718   104730: contig of 4013 bp in length
            *  104731   104830: gap of unknown length
            *  104831   115104: contig of 10274 bp in length
            *  115105   115204: gap of unknown length
            *  115205   156396: contig of 41192 bp in length
            *  156397   156496: gap of unknown length
            *  156497   179936: contig of 23440 bp in length
            *  179937   180036: gap of unknown length
            *  180037   186199: contig of 6163 bp in length.
FEATURES             Location/Qualifiers
     source          1. .186199
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBc0334A01"
                     /clone_lib="CHORI-201 Maize B73 BAC MboI Library"
     gap             100618. .100717
                     /estimated_length=unknown
     gap             104731. .104830
                     /estimated_length=unknown
     gap             115105. .115204
                     /estimated_length=unknown
     gap             156397. .156496
                     /estimated_length=unknown
     gap             179937. .180036
                     /estimated_length=unknown
ORIGIN

  Query Match             5.9%;  Score 77;  DB 14;  Length 186199;
  Best Local Similarity   71.4%;  Pred. No. 0.00023;
  Matches  115;  Conservative    0;  Mismatches   45;  Indels    1;  Gaps    1;

Qy        780 CTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTT 839
              ||||     ||  ||| || |||||||||| |||||| ||| ||  ||  |||||| ||||
Db     162058 CTTTTCTTTTACTTTTAGGGCTAGTTTGGGAACCCTATTT-TTCTAAGAAATTTTCATTT 162116

Qy        840 TTTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAG 899
              ||  ||   ||||||||| |||||||||| | ||||||||||    || | |||||||||
Db     162117 TTCTAAGTGAAATTAGTTCATTTTCTCTTAACAAAATAGAAATTTATTGGTAAAATAGAG 162176

Qy        900 TTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAA 940
              || ||| |||| ||||  | | | |    |||  |  ||||
Db     162177 TTTCCAAACTAACCCTCAAGTATATAATGAATGCAGGACAA 162217


RESULT 10
CR936840
LOCUS       CR936840              161616 bp    DNA     linear   HTG 26-FEB-2005
DEFINITION  Danio rerio clone DKEY-91O21, *** SEQUENCING IN PROGRESS ***, 13
            unordered pieces.
ACCESSION   CR936840
VERSION     CR936840.2  GI:60302473
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 161616)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 26, 2005 this sequence version replaced gi:60279457.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zK91O21
            -------------- Summary Statistics
```

```
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 156361 bases at least Q40
            Consensus quality: 157330 bases at least Q30
            Consensus quality: 158179 bases at least Q20
            Insert size: 160416; sum-of-contigs
            Insert size: 160313; 4.7% error; agarose-fp
            Quality coverage: 6.55x in Q20 bases; sum-of-contigs Quality
            coverage: 6.56x in Q20 bases; agarose-fp
            ----------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 13 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     3158: contig of 3158 bp in length
            *    3159     3258: gap of 100 bp.
            *    3259    11860: contig of 8602 bp in length
            *   11861    11960: gap of 100 bp
            *   11961    21800: contig of 9840 bp in length
            *   21801    21900: gap of 100 bp
            *   21901    45034: contig of 23134 bp in length
            *   45035    45134: gap of 100 bp
            *   45135    47371: contig of 2237 bp in length
            *   47372    47471: gap of 100 bp
            *   47472    66727: contig of 19256 bp in length
            *   66728    66827: gap of 100 bp
            *   66828    79067: contig of 12240 bp in length
            *   79068    79167: gap of 100 bp
            *   79168    97393: contig of 18226 bp in length
            *   97394    97493: gap of 100 bp
            *   97494   103542: contig of 6049 bp in length
            *  103543   103642: gap of 100 bp
            *  103643   112675: contig of 9033 bp in length
            *  112676   112775: gap of 100 bp
            *  112776   137298: contig of 24523 bp in length
            *  137299   137398: gap of 100 bp
            *  137399   158790: contig of 21392 bp in length
            *  158791   158890: gap of 100 bp
            *  158891   161616: contig of 2726 bp in length.
FEATURES             Location/Qualifiers
     source          1. .161616
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-91021"
                     /clone_lib="DanioKey"
     misc_feature    1. .3158
                     /note="assembly_fragment:00020
                     fragment_chain:1"
     misc_feature    3259. .11860
                     /note="assembly_fragment:00240
                     fragment_chain:1"
     misc_feature    11961. .21800
                     /note="assembly_fragment:00342
                     fragment_chain:1"
     misc_feature    21901. .45034
                     /note="assembly_fragment:01299
                     fragment_chain:1"
     misc_feature    45135. .47371
                     /note="assembly_fragment:00002
                     fragment_chain:1"
     misc_feature    47472. .66727
                     /note="assembly_fragment:00606
                     fragment_chain:1"
     misc_feature    66828. .79067
                     /note="assembly_fragment:00463
                     fragment_chain:1"
     misc_feature    79168. .97393
                     /note="assembly_fragment:00794
                     fragment_chain:1"
     misc_feature    97494. .103542
                     /note="assembly_fragment:00074
                     fragment_chain:1"
     misc_feature    103643. .112675
                     /note="assembly_fragment:00146
                     fragment_chain:1"
     misc_feature    112776. .137298
                     /note="assembly_fragment:01585
                     fragment_chain:1"
     misc_feature    137399. .158790
                     /note="assembly_fragment:01032
                     fragment_chain:1"
     misc_feature    158891. .161616
                     /note="assembly_fragment:00044.0"
ORIGIN

  Query Match             5.7%;  Score 74.4;  DB 14;  Length 161616;
  Best Local Similarity   47.9%;  Pred. No. 0.00071;
  Matches  184;  Conservative    0;  Mismatches  200;  Indels    0;  Gaps    0;

Qy        576 TTTACAAACATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTT 635
              ||   |||     || |   | |||  |     ||  |     |     |  |  |  |
Db      79191 TTATAAAATNNNNAANAAAAAAAAAAAAAATATATAAAATTTTTNNTANNTTNNATTTAAT 79250

Qy        636 TCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAG 695
              | |||||  || || ||  |||| || |   ||||   ||   | | ||   ||  ||||
Db      79251 TATTTTTTTTTTTTTTTTTTTTTTTTTTTATATTTTATTATTATATTTTATTTTTAATAT 79310

Qy        696 AGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCA 755
              | |||||  ||    |||  |||| |  | |||  | |  |||||  | |          |
Db      79311 TTTTATTTTTTTATAATTAAATAATAAAAATATTTATTTTTTTTTTTTTTTTTTAAAA 79370

Qy        756 TGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCT 815
              |   |   | |  |  |||  | ||| |  || |||| | |  ||  |    |    |
Db      79371 ATTAAAATTATTTTNTNAAAATTTTTTTTTNTTTATTTTTATATATATATATATATATAT 79430

Qy        816 GTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAA 875
              | | ||| |||.||  ||    ||||    |   |||  ||| |||| | ||| |   |
Db      79431 TTATATTTAAAAAAAAATTTTANTTTTTTTTATTTATTTTTTTTTTTTTTTTTTTNTA 79490

Qy        876 TAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAAT 935
              |  |||| |  || | | ||| |  |       ||    || |    | |    ||| ||
Db      79491 TNAAAAAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 79550

Qy        936 TACAATCACTGTGTATAATTATTT 959
                ||   | | ||  ||| |||
Db      79551 ATATATATATATATAATATTTTTT 79574


RESULT 11
AC117267
LOCUS       AC117267               25769 bp    DNA     linear   INV 21-FEB-2004
DEFINITION  Dictyostelium discoideum chromosome 2 map 5836255-5862024 strain
            AX4, complete sequence.
ACCESSION   AC117267 AC115597
VERSION     AC117267.2  GI:42733680
KEYWORDS    HTG.
SOURCE      Dictyostelium discoideum
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 25769)
  AUTHORS   Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P.,
            Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K.,
            Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
            Noegel,A.A.
  TITLE     Sequence and analysis of chromosome 2 of Dictyostelium discoideum
  JOURNAL   Nature 418 (6893), 79-85 (2002)
   PUBMED   12097910
  REMARK    The Dictyostelium Genome Sequencing Consortium
REFERENCE   2  (bases 1 to 25769)
  AUTHORS   Baumgart,C.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (09-APR-2002) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
REFERENCE   3  (bases 1 to 25769)
  AUTHORS   Baumgart,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2004) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
COMMENT     On or before Feb 21, 2004 this sequence version replaced
            gi:19570016, gi:20087114.
            CDS predictions from GeneID may contain errors. Further Information
            is available from IMB Jena, Department of Genome Analysis
            (http://genome.imb-jena.de/dictyostelium/)
            and the Univerity Colonge, Institute for Biochemistry I
            (http://www.uni-koeln.de/dictyostelium/project.shtml
            Funding
            Agency : Deutsche Forschungsgemeinschaft (DFG).
FEATURES            Location/Qualifiers
     source         1. .25769
                    /organism="Dictyostelium discoideum"
                    /mol_type="genomic DNA"
                    /strain="AX4"
                    /db_xref="taxon:44689"
                    /chromosome="2"
                    /map="5836255-5862024"
     CDS            complement(join(832. .1070,1237. .1349,1533. .1798,
                    1896. .2053,2231. .2360))
                    /note="GeneID exon scores (in order of location ranges):
                    23.68, 1.54, 15.75, 9.21, -4.75 - GSCJ_ID dd_03387"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="AAS38625.1"
                    /db_xref="GI:42733681"
                    /translation="MRLQIILFLSILLIFQLKSTFSSNDISSTILSITPNFNKDYNNQ
                    ILSKSEWFRSLEYSSNKKVIFEVFDGFNHEYIENLDRIIKEFRENKVYGVDFSFSFCL
                    NSKFNYLEVIKKFSTISRNNKVYPENIWLNIESTHENCYNNSNYLYNKNNNEYQEYLV
                    FSKQVIKSFLQNGYNIGIKSSKNTFIKIFENNKDKLFSSLPFFYLKNIAENQEFEDYY
                    SSPLGSWTDFKIKQFKKLNDIETILSDVEVNLCSSSNNNNNNNNVDDETLTLLGSGSG
                    SGSGSGSSGTSGAQLSGTGYVINHQ"
     CDS            join(2957. .3173,3272. .5583)
                    /note="GeneID exon scores (in order of location ranges):
                    24.88, 254.96 - GSCJ_ID dd_00753"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="AAS38626.1"
                    /db_xref="GI:42733682"
                    /translation="MEFPTKYVYITSNLDKPKEWFTQSEMSLTTTTDTIQKSFVNNSG
                    SSSSSKGFGEAVLLDILDTAGQEEYSAMRDQYVRTGDCFMIVFSIDSRSSFEEVSQLK
                    QHIERVKDRDDVPIIIVGNKVDLESRRQVSRIEADQLARSLRVPYIETSAKTRSNIEE
                    AFFTLVRHTPRNTVYKVVVMGGGGVGKSAIIIQFIQNHFVEEYDPTIEDSYRKQVTIS
                    GLPPIGGSLNKKSSSSSSSSSSSSGKTGLFNKIFSGKDKQPSPQQAASPSTIDRTGQI
                    STNRLEANVLSYSMSNLSKEVPLITGDCVYCQGCNVILSRFSNLVKTGDDSFTWKCEF
                    CKYSNSNILLEQGEIPNKDSVEYVLSSPSTSSTTDGSKREESIIIYCIDVSGSMGITT
                    EVPSLQSEWVNAKKGVKGASSGPSYISRLECVQSSIPTMIDRLSIQYPNKRVVLVTFS
                    DEVMIYTQSNSVDGPIVIAGDKLEDFDQLIEIGRSMTYDKLPTASGSSDFLKAKIKSL
                    EPVQSTALGPALLVSAAIASQKMLSEVVICTDGVPNVGLGAIEDLPLGPAQEFYEKVT
                    KLAQNNKTTINIIGISGSHIDLGVIGKVSEQTNGNITIIHPLELAREIRKLTQNPMIA
                    TDVEMSICLHPTLEINKYDSKQGLSRVVKQFPNVNSLTDLTLLYSSRNRPTEFVQIYP
                    FQIQIKYTKLDGVRCLRVVSAQLQATPDFNTSTSNANISILAMAFTQQAAKIAQQQEY
                    MESRLHLKAATKLIRSLCNTDEQWEEFYNFEVLREEMEAPLITCIKNKQQRVEKAATD
                    DEIQVFYKMKNVHKSFVEGGRKKDISRRKGEAEINKQYYNIKFT"
     CDS            complement(5711. .6796)
                    /note="GeneID exon scores (in order of location ranges):
                    68.61 - GSCJ_ID dd_00752"
                    /codon_start=1
                    /product="similar to Plasmodium falciparum (isolate 3D7).
                    Hypothetical protein"
                    /protein_id="AAS38627.1"
                    /db_xref="GI:42733683"
                    /translation="MSSRVKSLIERYNGINLTIEKSLSPRGLINTNINNNINNNNNNN
                    GTITTTTSGGNNNNNSTSPSTNPVSFDRVVDHSNNNQIQIVIDSFVNINNNNNNNNNN
                    NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNFNINQNKTNNQNQNQNHLNCN
                    KFNNNNNYYFDFSTNEMITTLNSCTLEQIGTIGNDCYSTNTTPSPTSSSSSTSSLLNN
                    KNSNHWMMPSLKINQIINDNNNNISNINFVQIQSNSSSLSSSCSTYQSCPIEDEHIQE
                    QLRIQKEQLKQIDDDDDDDDDDDDDDDSSPSPPPPPTTTTKTFCLSPRPIIKKEENID
                    QNSMSPPPKLDKRRFSRDLIPFTIANI"
     CDS            join(8924. .9076,9236. .10573)
                    /note="GeneID exon scores (in order of location ranges):
                    10.37, 128.60 - GSCJ_ID dd_00729"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="AAS38628.1"
                    /db_xref="GI:42733684"
                    /translation="MSSSTIITPIATTTTNNNSSTNGIHNSKTNRTLFKNIKNNIPNS
                    PNNAPIKKNSVGIIGNGPNVNNNNKDRKLNSNDGVNFGKKNILFTPSKNSMVSTTKHS
                    SSSSSLLAKNEYDIESIKKELKSMDISTGGISPLSSPLRESSPISTIDLESITATTTK
                    LTDAFVAASSSSSSVKYTIDVLLSFKSANTKRPIQIDIVENHQKEIVVPLSLPTTPYN
                    NNNNNNNNTNNSSQHNTANGQKYPIFSPQISPFKIAYAQSSSTKTNINNNNNNNNITP
                    TKKANSNITTPQSNNRNNHYNSNTKSSTKKQNPIPFSLNTATPNQKNTTTTPSKKSTT
                    TTPKSNKKINDVNAAFAAVANSTTTTEQPVAIIDGATPSTNESPKKTPLKETDPKRLA
                    ARQRQIDIGKMTAGYKNYIALVPKSKRKPTDPKTPNKNQVCSKRSWDGQIKKWRRQLH
                    ENYDWTGIPFNEKEVEASNNESTEESTLEKEMNKLEISEENENHNN"
     CDS            complement(10972. .11979)
                    /note="GeneID exon scores (in order of location ranges):
                    23.29 - GSCJ_ID dd_00728"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="AAS38629.1"
                    /db_xref="GI:42733685"
                    /translation="MEEKIKFYFEIIDFQNQKFKIQEFTSKLIGLKEESFTTFKPIVY
                    EKYLNWTQSIEESILKTNGTINKSIFEEIFSYCGYIGEFLEYEPPIVFLKFTTFGIIL
                    DDYIFEKINSLNMKLNEKEKLINSLIYYNNKNENKIGFEFWEIINEFQNFTHKESFQR
                    IINANSLWIKSSIDSRNISNSPINSRCSFNEYFEKRSSDASGDFILTISMIGIMDNYI
                    ENSIIESKEFQTINYHAKSFFLLINDLYSFNREINENDLLNYIKILAIQLNSIQLSID
                    KTIELIIDHYLKFLSSIETILKLYQNDQSTYQLLKQVFQNSNKILSGIYFAHKKSKRY
                    N"
     CDS            join(12884. .13020,13172. .15320)
                    /note="GeneID exon scores (in order of location ranges):
                    2.42, 136.56 - GSCJ_ID dd_00727"
                    /codon_start=1
                    /product="similar to Dictyostelium discoideum (Slime
                    mold). Hypothetical 97.7 kDa protein"
                    /protein_id="AAS38630.1"
                    /db_xref="GI:42733686"
                    /translation="MDNLIYDFLIKKEYNQQSFLDCILILSINGLYISYDGLLDYCNN
                    FRLIENKTVNKEAISYYGLNYNQYQLLDEFTRLLKLKNKTTDCKNKIINTHNSLNKCC
                    GNNKFDYSILIIIENNNNIINKDYQFNNNNKNSNNNIYNNNNDNNNSIIEEKDLNKII
                    TLKNERINGNDGEITTTTTTTTTSRSSTTIMTTTTTTNGVNDRNYLLFRQVWRNIVIK
                    TQILFHLRLYNIHANKKVFLTPIQLVDYKFKGYLQSMVLNYHDDNDDDADEDSYDSDI
                    DDDDSYSDGGNCSSSSSGSSDIGSSSNSINNGISNSSSSSILSNSSLLELSMIRVNSL
                    PEGLKSIEFEKEYNVIQDRLLPPSISSIRFSYGFNQRIAKGVISDNVISITFGDSFNQ
                    SLDGNWLPKQLKHLQFGHKFQQTIKMGQLPSSLTSLILDPRSYKGVIEIGSIPDSVKT
                    LDYKFNSCSNQESISFNPIPKSITRLVFDSEFNQNIKANDISSNYNLTSIHFGEHFNS
                    DIGIKSLPNSIREIKFGRAFDRDIKLCPSSITSIDFGNKFNRPLSMMTQTLTSIDFGS
                    KFNQIIPQGIFIHTKLKSLNFGYHFNQIIPADTLPPTLESLNLGGYNREITVKNDEYD
                    CYGISNKGGFGSNSSSNFCVGGTNNGLREMLKNTTSLKTLTLNYFNRKIEVGDLPNSI
                    ESLNLGYHFNQPIGNNVLPKLLKKLFILNSEFNQNISADGCIPFGLQTIYIRNSNMNF
                    INSLNPLFITKYINIIDLSHL"
     CDS            join(16184. .16199,17272. .17613,17707. .19331)
                    /note="GeneID exon scores (in order of location ranges):
                    0.08, 32.32, 227.42 - GSCJ_ID dd_00726"
                    /codon_start=1
                    /product="similar to Dictyostelium discoideum (Slime
                    mold). Spore coat protein SP87 precursor (PL3 protein)"
                    /protein_id="AAS38631.1"
                    /db_xref="GI:42733687"
                    /translation="MFIGEDCNKITNEEECHKSSECIVINYTPCCGEQKWACSKGTFD
                    TCTYENSCYRNSSNNQVVEVSNKCFNLDGFIKITTPTEYSCSDAKIKECALLGKSCSF
                    QKNSCSNPTSCCPGESICEGLSSGSSTSGGGSSGGTSGGSSSGGTSGGSSSGGTSGSS
                    SSGSSSGGVSSCSTTHCPEGYHCSMVNDVATCLASTTGGTGLPGTSSSTAGVSSCLTT
                    LCPIGHICVEDSNGVNCVPNGGGTSGGSSSTGTSGGHPDPCRDVDCPDGFHCECKDGK
                    TAKCVPSPTTGSSSTSGGHPDPCKDVTCPDGFHCECKDGKTAKCVPSPTTGSSSTSGG
                    NTNPCSNVNCPDGFYCECKDGKTAKCVPSGPTQPPKPPVCSLRCPPNHECRFNDQGHQ
                    CCVKVHHDRCSLRCPHGHECKVDHNGKECCVRSHRPPPPEVCSLRCPPKHECKFDDHG
                    KKCCVKIHCDEVCDLDCGRGFECKIRHDGSKCCVRSERPHPPQHEKCNKRCPPGHECK
                    VDQHGKECCVVAHRPPPKCSLRCPPRHECRVNHFGEECCVKVHHDKCSLRCPPGHECK
                    VDQHGKECCVVAHRPPPKCSLRCPPKHECRINHFGEECCVKSRNDCLTCEDLNCERKG
```

```
                     LHCAMKTVPIDKENCCEKVPVCYSNNPLLDGGHGFI"
     CDS             19762. .21336
                     /note="GeneID exon scores (in order of location ranges):
                     106.68 - GSCJ_ID dd_00725"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAS38632.1"
                     /db_xref="GI:42733688"
                     /translation="MDNNNLPQPYFPQSNNYCDCDDNENGNDSTLTIQPKPIAPSLPS
                     SQQPSPTIISQINHNKYQLFKEFKEHNNNIDSKMLCVQLSEFQISPKTATNTFDSWLN
                     SLWLPPTGFKKLVDNHLKKVYIPFYSYRVNYNVEFSGFVTYDSKLNTNALYDKYGNTN
                     STYNNNNYYNVNNNSLNNSNNADKNKKTITSTSATILNKENVELYNNGNNNNLNNKNS
                     EDILKDKSIKWNKVDGYITRTINRILLLACNSMDLLSDESIMNCMVSQQFDFNNVIEK
                     SLDEFNQFAQLDQVEVLARYEPDVVWKRDYLNHLLETEQSMSIEYLKQQEQSKFVKNV
                     SSTIQFNEKLFSLVYIPFYYTKYEFQNQSFSFFLNSQNGVISATRPEIGLGKIGDFLK
                     YTQNYFSTLVGNFEPFSKCKGDELAIIDSFKVYNTSSYFLTFSRSSTNFLGSSSIGHL
                     TIRNNSKSITATLRGQKRRSVQKGNPFSLLPLHERTFSFKGHWCFEIIDGEVNQIDVI
                     KINPREGGNIYQSIKN"
     CDS             complement(21961. .24357)
                     /note="GeneID exon scores (in order of location ranges):
                     173.48 - GSCJ_ID dd_00724"
                     /codon_start=1
                     /product="similar to Arabidopsis thaliana (Mouse-ear
                     cress). Putative PREG1-like negative regulator"
                     /protein_id="AAS38633.1"
                     /db_xref="GI:42733689"
                     /translation="MSKVYSKSQTVSPGKPQNIFVTALPSSVSAPSLLSLQKPNNGAG
                     SHRKQRQNSGSSKNRPNLPLDQIKSTERKNSGNKSNNNNNKNRKYSKNNSNVNSNNIN
                     SSSIVIKCEEQQQQQQQQQQQQQQQQQQQQQQQQQQQDGFLKSQEFPNLFKNKLYLKDDIS
                     SPSSSSSTSSPSCLSSSSSSTFSQSEIAENKDDINNNNCNNTNGNFEEEEEEEEKNKV
                     EQQSIEEETFIESSFEETKEVVEEVPFEEQFEILPISIGEETLKEYTSFLSVLSELIV
                     DMYNDDDKSSTSSQSSSSSSSSSAPSSPKPIKNNDNNTTTTTTTATTTTSKTATSS
                     LSTIEKFQVFNVDQIPEISIEAYIERVFKYLPFGTDIFIFSTIYLDRLIQWNQEIQIS

  Query Match             5.6%;  Score 73.4;  DB 2;  Length 25769;
  Best Local Similarity   49.4%;  Pred. No. 0.0014;
  Matches  217;  Conservative    0;  Mismatches  221;  Indels    1;  Gaps    1;

Qy       438 CCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGAAAGCTATCATGTAA 497
            | | | | || | ||||||| ||      ||      |  |   |||   |  |  ||
Db      2392 CGAAACAACGACTAAAAAAAAAAAAAATTTGTTGAGGAAAAAAAAAAAAAAAAAAAAAAA 2451

Qy       498 TGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTTCGAAAG 557
            | |  ||| ||| |   || ||||  | | |  |  ||||||. ||  ||||  |  |||
Db      2452 TATGAAAAAAAAAAAAAATAAAAAAAAAGGTTTTTACAAAAAAAAAAAAATATCTAAAAAA 2511

Qy       558 AGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAAGGTAAG 617
            | | | ||||   |   |||| |   || |||  |  |  ||| | |    ||  |||
Db      2512 AAATTATTTATTAAATTATTTAAATTGATTTAATTGTTAACAATAAAAAAACAAAATAAA 2571

Qy       618 CTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACAT 677
              ||  |  || |  |  || ||     ||| || |  | || ||||| ||  |  | | ||
Db      2572 ATAAAATATTAATTTAATTATTAAA-ATTATTTTCTTATTTTTATTTTTAATTGAAAAAT 2630

Qy       678 TTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATT 737
             |   || |   | |       ||  |||| |||   |||| |    | |   | |  | ||
Db      2631 TAATTTTTAAAAGATATTTGTTTTTTTTTTCTTTTTTTTTTTTTTTTAAAGTGATCTATTT 2690

Qy       738 TTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAG 797
             || | |   | |         |   | ||  || || |||  | ||   ||  | | |
Db      2691 TTAAAAATTGGGTGTTTGATTGAAAAATATTCTTTTAAATTTTTTCAGGAAATATCTTTT 2750

Qy       798 AGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTTCAAAAAAAATTAGTT 857
               |  |||   ||| |       |||       ||||  ||||||  ||||||  | | ||
Db      2751 TTTTTTTTTTCAAACACCACCAATTTTTTTTTTTTTTTTTTTTTTTAAAAAACCTAAATT 2810

Qy       858 TATTTTCTCTTTATAAAAT 876
             |  ||| | ||  |   ||
Db      2811 TTATTTTTTTTATTTTTAT 2829

RESULT 12
AC155379/c
```

```
LOCUS       AC155379              172293 bp    DNA     linear   HTG 25-JAN-2005
DEFINITION  Zea mays strain B73 clone ZMMBBb0131C15, *** SEQUENCING IN PROGRESS
            ***, 27 unordered pieces.
ACCESSION   AC155379
VERSION     AC155379.2  GI:58082241
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 172293)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Consortium for Maize Genomics - BAC skim sequencing and assembly
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 172293)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
REFERENCE   3  (bases 1 to 172293)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
COMMENT     On Jan 25, 2005 this sequence version replaced gi:57862900.
            -------------- Trace submission
                Center_name: TIGR
                Seq_lib_id: ZOCF
            -------------- Project informaton
                Web site: http://www.tigr.org/tdb/tgi/maize/
                Contact: maize@tigr.org
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 27 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2421: contig of 2421 bp in length
            *     2422     2521: gap of unknown length
            *     2522     5334: contig of 2813 bp in length
            *     5335     5434: gap of unknown length
            *     5435     8098: contig of 2664 bp in length
            *     8099     8198: gap of unknown length
            *     8199    10614: contig of 2416 bp in length
            *    10615    10714: gap of unknown length
            *    10715    12841: contig of 2127 bp in length
            *    12842    12941: gap of unknown length
            *    12942    15185: contig of 2244 bp in length
            *    15186    15285: gap of unknown length
            *    15286    36327: contig of 21042 bp in length
            *    36328    36427: gap of unknown length
            *    36428    44529: contig of 8102 bp in length
            *    44530    44629: gap of unknown length
            *    44630    58058: contig of 13429 bp in length
            *    58059    58158: gap of unknown length
            *    58159    66276: contig of 8118 bp in length
            *    66277    66376: gap of unknown length
```

```
            *    66377    70448: contig of 4072 bp in length
            *    70449    70548: gap of unknown length
            *    70549    74167: contig of 3619 bp in length
            *    74168    74267: gap of unknown length
            *    74268    82175: contig of 7908 bp in length
            *    82176    82275: gap of unknown length
            *    82276    83496: contig of 1221 bp in length
            *    83497    83596: gap of unknown length
            *    83597   114544: contig of 30948 bp in length
            *   114545   114644: gap of unknown length
            *   114645   116257: contig of 1613 bp in length
            *   116258   116357: gap of unknown length
            *   116358   119322: contig of 2965 bp in length
            *   119323   119422: gap of unknown length
            *   119423   121368: contig of 1946 bp in length
            *   121369   121468: gap of unknown length
            *   121469   125359: contig of 3891 bp in length
            *   125360   125459: gap of unknown length
            *   125460   126905: contig of 1446 bp in length
            *   126906   127005: gap of unknown length
            *   127006   131580: contig of 4575 bp in length
            *   131581   131680: gap of unknown length
            *   131681   134564: contig of 2884 bp in length
            *   134565   134664: gap of unknown length
            *   134665   138587: contig of 3923 bp in length
            *   138588   138687: gap of unknown length
            *   138688   156638: contig of 17951 bp in length
            *   156639   156738: gap of unknown length
            *   156739   167111: contig of 10373 bp in length
            *   167112   167211: gap of unknown length
            *   167212   168221: contig of 1010 bp in length
            *   168222   168321: gap of unknown length
            *   168322   172293: contig of 3972 bp in length.
FEATURES             Location/Qualifiers
     source          1. .172293
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0131C15"
     gap             2422. .2521
                     /estimated_length=unknown
     gap             5335. .5434
                     /estimated_length=unknown
     gap             8099. .8198
                     /estimated_length=unknown
     gap             10615. .10714
                     /estimated_length=unknown
     gap             12842. .12941
                     /estimated_length=unknown
     gap             15186. .15285
                     /estimated_length=unknown
     gap             36328. .36427
                     /estimated_length=unknown
     gap             44530. .44629
                     /estimated_length=unknown
     gap             58059. .58158
                     /estimated_length=unknown
     gap             66277. .66376
                     /estimated_length=unknown
     gap             70449. .70548
                     /estimated_length=unknown
     gap             74168. .74267
                     /estimated_length=unknown
     gap             82176. .82275
                     /estimated_length=unknown
     gap             83497. .83596
                     /estimated_length=unknown
     gap             114545. .114644
                     /estimated_length=unknown
     gap             116258. .116357
                     /estimated_length=unknown
     gap             119323. .119422
                     /estimated_length=unknown
     gap             121369. .121468
                     /estimated_length=unknown
     gap             125360. .125459
                     /estimated_length=unknown
     gap             126906. .127005
                     /estimated_length=unknown
     gap             131581. .131680
                     /estimated_length=unknown
     gap             134565. .134664
                     /estimated_length=unknown
     gap             138588. .138687
                     /estimated_length=unknown
     gap             156639. .156738
                     /estimated_length=unknown
     gap             167112. .167211
                     /estimated_length=unknown
     gap             168222. .168321
                     /estimated_length=unknown
ORIGIN

  Query Match             5.5%;  Score 72.4;  DB 14;  Length 172293;
  Best Local Similarity   78.6%;  Pred. No. 0.0016;
  Matches   99;  Conservative    0;  Mismatches   26;  Indels    1;  Gaps    1;

Qy        791 TTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAA 850
              | ||||||||||||||||||||| | |||| || ||| ||||||  |||||  |||  |||
Db      27185 TGTTAAGAGCTAGTTTGGCAACTCCGTTT-TTCCAAGGAATTTCTATTTTCCCAAGGAAA 27127

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
              ||||||| |||| |||||   |||||   || |   ||||||||| ||||| ||| ||||
Db      27126 ATTAGTTCATTTACTCTTGGGAAAATGAGAATCCTCTAGAAAAATGGAGTTTCCAAACTA 27067

Qy        911 GCCCTA 916
              ||||||
Db      27066 GCCCTA 27061


RESULT 13
AC155618/c
LOCUS       AC155618              201985 bp    DNA     linear   HTG 25-JAN-2005
DEFINITION  Zea mays strain B73 clone ZMMBBc0344D04, *** SEQUENCING IN PROGRESS
            ***, 21 unordered pieces.
ACCESSION   AC155618
VERSION     AC155618.2  GI:58082477
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 201985)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Consortium for Maize Genomics - BAC skim sequencing and assembly
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 201985)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
REFERENCE   3  (bases 1 to 201985)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
```

```
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
COMMENT     On Jan 25, 2005 this sequence version replaced gi:57863139.
            -------------- Trace submission
                Center_name: TIGR
                Seq_lib_id: ZGGW
            -------------- Project informaton
                Web site: http://www.tigr.org/tdb/tgi/maize/
                Contact: maize@tigr.org
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 21 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1916: contig of 1916 bp in length
            *    1917     2016: gap of unknown length
            *    2017    22480: contig of 20464 bp in length
            *   22481    22580: gap of unknown length
            *   22581    30468: contig of 7888 bp in length
            *   30469    30568: gap of unknown length
            *   30569    65218: contig of 34650 bp in length
            *   65219    65318: gap of unknown length
            *   65319    78774: contig of 13456 bp in length
            *   78775    78874: gap of unknown length
            *   78875    87797: contig of 8923 bp in length
            *   87798    87897: gap of unknown length
            *   87898    94463: contig of 6566 bp in length
            *   94464    94563: gap of unknown length
            *   94564   121761: contig of 27198 bp in length
            *  121762   121861: gap of unknown length
            *  121862   124636: contig of 2775 bp in length
            *  124637   124736: gap of unknown length
            *  124737   131029: contig of 6293 bp in length
            *  131030   131129: gap of unknown length
            *  131130   138739: contig of 7610 bp in length
            *  138740   138839: gap of unknown length
            *  138840   146859: contig of 8020 bp in length
            *  146860   146959: gap of unknown length
            *  146960   152969: contig of 6010 bp in length
            *  152970   153069: gap of unknown length
            *  153070   155026: contig of 1957 bp in length
            *  155027   155126: gap of unknown length
            *  155127   156434: contig of 1308 bp in length
            *  156435   156534: gap of unknown length
            *  156535   173224: contig of 16690 bp in length
            *  173225   173324: gap of unknown length
            *  173325   174858: contig of 1534 bp in length
            *  174859   174958: gap of unknown length
            *  174959   182700: contig of 7742 bp in length
            *  182701   182800: gap of unknown length
            *  182801   185872: contig of 3072 bp in length
            *  185873   185972: gap of unknown length
            *  185973   199501: contig of 13529 bp in length
            *  199502   199601: gap of unknown length
            *  199602   201985: contig of 2384 bp in length.
FEATURES             Location/Qualifiers
     source          1. .201985
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBc0344D04"
     gap             1917. .2016
                     /estimated_length=unknown
     gap             22481. .22580
                     /estimated_length=unknown
     gap             30469. .30568
                     /estimated_length=unknown
     gap             65219. .65318
                     /estimated_length=unknown
     gap             78775. .78874
                     /estimated_length=unknown
     gap             87798. .87897
                     /estimated_length=unknown
     gap             94464. .94563
                     /estimated_length=unknown
     gap             121762. .121861
                     /estimated_length=unknown
     gap             124637. .124736
                     /estimated_length=unknown
     gap             131030. .131129
                     /estimated_length=unknown
     gap             138740. .138839
                     /estimated_length=unknown
     gap             146860. .146959
                     /estimated_length=unknown
     gap             152970. .153069
                     /estimated_length=unknown
     gap             155027. .155126
                     /estimated_length=unknown
     gap             156435. .156534
                     /estimated_length=unknown
     gap             173225. .173324
                     /estimated_length=unknown
     gap             174859. .174958
                     /estimated_length=unknown
     gap             182701. .182800
                     /estimated_length=unknown
     gap             185873. .185972
                     /estimated_length=unknown
     gap             199502. .199601
                     /estimated_length=unknown
ORIGIN

  Query Match          5.5%;  Score 71.6;  DB 14;  Length 201985;
  Best Local Similarity   76.9%;  Pred. No. 0.0022;
  Matches  100;  Conservative    0;  Mismatches   29;  Indels    1;  Gaps    1;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              | |||| |||||||||| |||| | ||| || ||| | ||||| |||||  |||| ||||
Db      44986 TATAAGGGCTAGTTTGGGAACCATATTT-TTCCAAGGGATTTTCATTTTCCCAAAGAAAA 44928

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              |||||| |||||| | |   |||||||||| | | | | |||||  |||| ||| |||||
Db      44927 TTAGTTCATTTTCCCATAGGAAAATAGAAATCCCATGGGAAAATGTAGTTCCCAAACTAG 44868

Qy        912 CCCTAGAATG 921
              ||||| || |
Db      44867 CCCTAAAAAG 44858


RESULT 14
AF034389/c
LOCUS       AF034389               713 bp    DNA     linear   INV 04-FEB-1999
DEFINITION  Plasmodium falciparum sexual stage antigen (s16) gene, promoter and
            partial cds.
ACCESSION   AF034389
VERSION     AF034389.1  GI:3098290
KEYWORDS
SOURCE      Plasmodium falciparum (malaria parasite P. falciparum)
  ORGANISM  Plasmodium falciparum
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 713)
  AUTHORS   Dechering,K.J., Kaan,A.M., Mbacham,W., Wirth,D.F., Eling,W.,
            Konings,R.N. and Stunnenberg,H.G.
  TITLE     Isolation and functional characterization of two distinct
            sexual-stage-specific promoters of the human malaria parasite
            Plasmodium falciparum
  JOURNAL   Mol. Cell. Biol. 19 (2), 967-978 (1999)
   PUBMED   9891033
```

```
REFERENCE   2  (bases 1 to 713)
  AUTHORS   Dechering,K.J., Kaan,A.M. and Konings,R.N.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-1997) Molecular Biology, University of Nijmegen,
            Toernooiveld 1, Nijmegen 6525 ED, The Netherlands
FEATURES             Location/Qualifiers
     source          1. .713
                     /organism="Plasmodium falciparum"
                     /mol_type="genomic DNA"
                     /strain="NF54"
                     /db_xref="taxon:5833"
     gene            1. .>713
                     /gene="s16"
     promoter        1. .536
                     /gene="s16"
     mRNA            537. .>713
                     /gene="s16"
     5'UTR           537. .710
                     /gene="s16"
     CDS             711. .>713
                     /gene="s16"
                     /note="Pfs16"
                     /codon_start=1
                     /product="sexual stage antigen"
                     /protein_id="AAD12581.1"
                     /db_xref="GI:3098291"
                     /translation="M"
ORIGIN

  Query Match             5.4%;  Score 71.2;  DB 2;  Length 713;
  Best Local Similarity   45.8%;  Pred. No. 0.0063;
  Matches  282;  Conservative    0;  Mismatches  333;  Indels    1;  Gaps    1;

Qy        444 CAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGAAAGCTATCATGTAATGTTTA 503
              ||| |   ||||||||||  ||   ||  ||| | |  |||   |  ||  |||||
Db        681 CAAAAGAAAAAAAAATAAAAAATAAAAAAAATAAAATAAAAAAAAAAAATAAAATGTAAG 622

Qy        504 AAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTTCGAAAGAGACAA 563
              || |     |||| | ||  | | |  | |  |||     |   |||    ||    | |
Db        621 AATATTATCCTATGATAATAAAGCTATTGTGGAAATTATATTTTATGGAAAATATAATAT 562

Qy        564 ATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATGA 623
              ||     || |  ||       | | | |   | ||  |||||  | |  ||   | |
Db        561 ATATATGTAAATATTTTTTTTTTATTATATATA-AATATATATATATATATATTTTTT 503

Qy        624 CGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTT 683
                ||    |||||  ||||  ||  | |||||| || || ||||| |||| || |
Db        502 TTTTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTTTTTTTTTTATAAATATAAAA 443

Qy        684 CTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATA 743
              | | |  ||| |  |  |||       | | | || ||||   |  |||   ||  |||
Db        442 TTTTCAGTATAAAAAGTATTTAACGGAAATATGTAGAATGGAAAAGAAGAATTTGGTATG 383

Qy        744 TAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAG 803
              |||      ||    || |      |    | ||| ||| |     ||   | |   ||
Db        382 TAAATTTTAATATAAAATATAAAAAAGAAATATAATTTTTTTATAAATTAATATAAAAAG 323

Qy        804 TTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTT 863
              |||    |   |  |   |  ||    |||| |  |||| ||   |  ||   |||| |||
Db        322 TTTATTTATTTTTATATATATAATTTTTTTTTAAATTTAAATTCATATATTTTTAATTT 263

Qy        864 CTCTTTATAAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTT 923
              | ||| |    |         |||| | | |||   |     |     |        |  ||
Db        262 ATATTTTTTTTTTTTTTTTTGCTTATATATATAATATATACGTAATATAATATAATTT 203

Qy        924 TTCCCAATAAATTACAATCACTGTGTATAATTATTTGGCCAGCCCCATAAATTATTTAAA 983
              ||    |  ||  |         |||| ||||||      |  |||| || | | | |
Db        202 TTTTTTTTTTTTTTTTAATTTGAAATATATTTATTTTATTTCTACATATATAAATAATA 143

Qy        984 CCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGA 1043
                ||  | | ||     |        ||    | | ||| |  |    |  ||||| |
Db        142 TAAAATTTTATATTTTAAGGTGTATCATTAATGTTTTTTTTAAAATAAAAAAAAAATAGT 83

Qy       1044 GAGAGAGAGGAAGAAA 1059
                |   | |   | |||
Db         82 TATCTATATATAAAAA 67


RESULT 15
DDU86962/c
LOCUS       DDU86962                3576 bp    DNA     linear   INV 23-JUL-2001
DEFINITION  Dictyostelium discoideum MigA (migA) gene, complete cds.
ACCESSION   U86962
VERSION     U86962.1  GI:1841871
KEYWORDS    .
SOURCE      Dictyostelium discoideum
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 3576)
  AUTHORS   Escalante,R., Wessels,D., Soll,D.R. and Loomis,W.F.
  TITLE     Chemotaxis to cAMP and slug migration in Dictyostelium both depend
            on migA, a BTB protein
  JOURNAL   Mol. Biol. Cell 8 (9), 1763-1775 (1997)
   PUBMED   9307972
REFERENCE   2  (bases 1 to 3576)
  AUTHORS   Escalante,R., Wessels,D., Soll,D. and Loomis,W.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1997) Department of Biology, University of
            California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-032,
            USA
FEATURES             Location/Qualifiers
     source          1. .3576
                     /organism="Dictyostelium discoideum"
                     /mol_type="genomic DNA"
                     /strain="AX4"
                     /sub_strain="AK244"
                     /db_xref="taxon:44689"
                     /chromosome="4"
                     /map="4.25"
     gene            884. .3423
                     /gene="migA"
     mRNA            join(<884. .2899,2998. .>3423)
                     /gene="migA"
                     /product="MigA"
     CDS             join(884. .2899,2998. .3423)
                     /gene="migA"
                     /function="essential for slug migration"
                     /note="N-terminus of this protein is similar to other BTB
                     domain proteins, such as the Mus musculus and human zinc
                     finger proteins encoded by GenBank Accession Numbers
                     U14556 and L16896, respectively"
                     /codon_start=1
                     /product="MigA"
                     /protein_id="AAB47544.1"
                     /db_xref="GI:1841872"
                     /translation="MATINIINNNNNQNNNNQNNNNNINHTEYFNNGFKSLLGEFKKY
                     TDIKIKCGGQIFPTHRIILANSSEFFGRLLLSDFKESSQSVIELKQPDPFNVFPMVLE
                     FMYDGIIYLSPENVIPLLAMSDHYLIRDLQHLCLLFLDQNLHRENVLTILKYSIDFHF
                     EPITERCITIISKNFTFVCENLYKQTKINQNNNNNCINIENNKKHLNTSSASSSLNSS
                     TSSASSSSSSSNNNYYTTTTTTTTTTTASSSNFNKNLTNCKVTGQQQQQQQQQQQQQQ
                     QQQQQQQQQQPIYDFSFLPASIFYQILDHDYLVVNSEYTLYRIVLNFIEKFFGENFTE
                     ITYNTNITDQLKVSDLMSQIRYILMPYDQLTLITKDPLIPKEILFETLLERLKRFEGK
                     TSSPSSSSTSSSPSNTSNSPTKQTSPSKYQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
                     PQQQPQQQPQQPQPQIIQQQPQQPQIIQQQPQQLQQLQQPQQPQTIQQMVQQQQNIQQ
                     QQQQQQQQQNIQQQQTQNMQQQLQQQLLQQIQQQQQQQQQPQIIKQFTPSSSPLSSSF
                     ISLNASQSKLSINNATNTTNNNNNNSNNVTNTNNNNIINNNNNNNNKQTQSNNNDNKI
                     LKRYQPRPPQSILFEYSFDFDFKGIIFWISTDGGNEKWSNPHSTSKIKITSSSIDKGN
                     LYDIVELTPNAFWTKDVPASWVMIDLGPNRTVVPMYYTIRHGLSYKSDSLRTWDFQGS
                     TNGEQWTVLKRHTNDPSLNYKYATHSWPVTGCETAFRYFRILQTGKNSNNRNFLVIGG
                     LEIYGELCETNPNPN"
     exon            <884. .2899
                     /gene="migA"
                     /number=1
     misc_feature    951. .1289
                     /gene="migA"
                     /note="encodes BTB domain; similar to the Drosophila
```

melanogaster BTB domain encoded by GenBank Accession Number U01333"
variation 1253^1254
/gene="migA"
/note="gene interruption in mutant"
/replace="pbsr1 external plasmid, approximately 4 kb"
intron 2900. .2997
/gene="migA"
/number=1
exon 2998. .>3423
/gene="migA"
/number=2

ORIGIN

Query Match 5.4%; Score 71.2; DB 2; Length 3576;
Best Local Similarity 47.2%; Pred. No. 0.0049;
Matches 217; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

```
Qy        412 TGTCTAGTGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAAC 471
              | | || | |    ||||   || |    |     |  | ||||||  ||  | |  ||
Db        852 TATTTATTATTATTTATTATTACAATTTTGTTTTTGTATAAAAAAATAAAAAATAAAAAA 793

Qy        472 TAATCCACTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCT 531
              |||   |    |||   |  ||  ||  |  |||  || | | |  |||    |   |
Db        792 TAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAACAAACAAATTAAATTTTAA 733

Qy        532 CTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAG 591
              ||||||||  |  | ||    ||  | | ||| ||   || |   ||  ||    ||
Db        732 TTTAAAAAAATAAAAAATAAAAAATAAAATAAAATAAAATAAAAAATAATAATTAAAAAA 673

Qy        592 AGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGT 651
              |   | |||  | |   |||  || | | ||       |   |||| |||||  ||  |
Db        672 AAAAAAAAAAAAAAAATAAATATAGGGTGTGTGAAGAATTGTTTTTTTTTTTGTTTGTTG 613

Qy        652 TATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGA 711
              |  ||  || |  | ||| || | | ||| ||      |  | | |     ||      |
Db        612 TTGTTATTTTTGTTATTTTTAAATAATTTATTTATACAATAAAAATACTTTAAACACAAC 553

Qy        712 TTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGT 771
              |  | |  || ||     |  ||||| | |       |  | | |  ||   | |
Db        552 ACACAACACACCACAATTGTTTTTTTTTTGTGTTTTTTTGTGTTTTTTTTTTATCAACCT 493

Qy        772 TCAAAAATCTTTCTGATTTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAAT 831
              |  ||   | || ||||||  ||      | |||    |     ||| |||  || |||
Db        492 TTCAATCGCCTTTTGATTTAGTTTTTTTTTTTTTTTTTTTGAATTGTTTTTTTTTGAAAAAT 433

Qy        832 TTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTAT 871
              ||  |||| |  ||   |||  |||||||| | ||| |
Db        432 ATTTTTTTTATTTAATTTTATTTTTTTATTTTATTTTTTT 393
```

Search completed: March 5, 2006, 21:55:43
Job time : 5723.47 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:08:23 ; Search time 872.437 Seconds
(without alignments)
10014.946 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311
Perfect score: 1311
Sequence: 1 ccatggtgtctctatgaaaa.....cgcacatatgtttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1311	100.0	1394	2	AAX07408	Aax07408 Zea mays
	2	1311	100.0	1394	2	AAX07409	Aax07409 Zea mays
	3	1311	100.0	1394	5	AAH76332	Aah76332 Z. mays M
	4	1311	100.0	1394	5	AAH76333	Aah76333 Z. mays M
	5	160.6	12.3	255	5	AAH76340	Aah76340 Z. mays M
	6	146	11.1	158	5	AAH76334	Aah76334 Z. mays M
	7	74	5.6	320	13	ADX48600	Adx48600 Plant ful
	8	74	5.6	624	13	ADX33444	Adx33444 Plant ful
	9	72.8	5.6	2445	13	ADX61650	Adx61650 Plant ful
c	10	71.8	5.5	928	13	ADX34996	Adx34996 Plant ful
	11	71.8	5.5	2537	13	ADX61114	Adx61114 Plant ful
	12	71.8	5.5	3607	13	ADX60218	Adx60218 Plant ful
c	13	70.2	5.4	883	4	AAL15210	Aal15210 Human bre
c	14	70.2	5.4	960	11	ACN85231	Acn85231 Breast ca
	15	69.8	5.3	13400	13	ADR04296	Adr04296 Corn FT h
c	16	69.2	5.3	527	13	ADX09466	Adx09466 Plant ful
c	17	68.4	5.2	1215	13	ADX49916	Adx49916 Plant ful
	18	68	5.2	1092	14	AEB28609	Aeb28609 Maize pol
	19	67.8	5.2	928	13	ADX34996	Adx34996 Plant ful
c	20	67.4	5.1	1260	13	ADX60038	Adx60038 Plant ful
	21	67.2	5.1	2232	13	ADX33787	Adx33787 Plant ful
	22	67.2	5.1	2249	13	ADX45524	Adx45524 Plant ful
	23	67.2	5.1	2657	2	AAZ10551	Aaz10551 DNA seque
c	24	66.4	5.1	6027	2	AAX58751	Aax58751 Maize dul
c	25	66.4	5.1	6027	8	ABX09935	Abx09935 DNA encod
c	26	66.4	5.1	6027	12	ADK12106	Adk12106 cDNA enco
	27	66.2	5.0	439	8	ABX35844	Abx35844 Bovine ES
	28	66	5.0	1376	13	ADO83353	Ado83353 Plant ful
c	29	65.6	5.0	346	4	AAI87279	Aai87279 Human pol
	30	65.6	5.0	2274	13	ADO82110	Ado82110 Plant ful
	31	65.6	5.0	2274	13	ADO82057	Ado82057 Plant ful
c	32	65.6	5.0	8056	8	ABZ10246	Abz10246 Haematopo
c	33	64.6	4.9	9859	13	ADS89714	Ads89714 Oligonucl
c	34	64.4	4.9	9859	13	ADS89440	Ads89440 Oligonucl
c	35	64	4.9	8056	8	ABZ10100	Abz10100 Haematopo
	36	63.8	4.9	431	8	ABX44556	Abx44556 Bovine ES
c	37	63.6	4.9	335913	5	AAI61371	Aai61371 Soybean 2
c	38	63.6	4.9	335913	5	AAI61372	Aai61372 Soybean 2
	39	63.4	4.8	12968	4	AAS45494	Aas45494 Chemical1
	40	63.4	4.8	12968	4	AAS46779	Aas46779 Tumour su
	41	63.4	4.8	12968	6	ABL34085	Abl34085 Human imm
	42	63.4	4.8	12968	6	ABK28424	Abk28424 DNA trans
	43	63.4	4.8	12968	6	ABN80285	Abn80285 Human che
	44	63.2	4.8	1340	14	ADZ71035	Adz71035 Human chr
	45	63	4.8	7624	6	ABL34113	Abl34113 Human imm

ALIGNMENTS

RESULT 1
AAX07408
ID AAX07408 standard; DNA; 1394 BP.
XX
AC AAX07408;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW Ms45; male; tissue-preferred; regulatory region; plant cells;
KW plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
PN WO9859061-A1.
XX
PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98WO-US012895.
XX
PR 23-JUN-1997; 97US-00880499.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
DR WPI; 1999-105628/09.
XX
PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
PS Claim 2; Page 22-23; 39pp; English.
XX
CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue- preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile

```
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 1311;  DB 2;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-269;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 2
AAX07409
ID   AAX07409 standard; DNA; 1394 BP.
XX
AC   AAX07409;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW   Ms45; male; tissue-preferred; regulatory region; plant cells;
KW   plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS   Zea mays.
XX
PN   WO9859061-A1.
XX
PD   30-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US012895.
XX
PR   23-JUN-1997;   97US-00880499.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman GA,  Kendall TL;
XX
DR   WPI; 1999-105628/09.
XX
PT   New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT   - useful in mediating plant fertility, especially hybrid seed production.
XX
PS   Claim 3; Page 23-24; 39pp; English.
XX
CC   The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC   region. It may be used in the construction of a vector for a method of
CC   producing exogenous genes in a male tissue- preferred manner, which is
CC   useful in restoring or conferring fertility, such as in hybrid seed
CC   production. In conferring fertility, a monocot/dicot plant is transformed
CC   with the exogenous nucleotide sequence (a male sterility gene, preferably
CC   Ms45), which encodes a product selected from auxins, rolB and diptheria
CC   toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 1311;  DB 2;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-269;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
Db       1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy      61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
Db      61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy     121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
Db     121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy     181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
Db     181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy     241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
Db     241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy     301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
Db     301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy     361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
Db     361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy     421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
Db     421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy     481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
Db     481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy     541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
Db     541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy     601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
Db     601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy     661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
Db     661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy     721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
Db     721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy     781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
Db     781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy     841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
Db     841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy     901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
Db     901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy     961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
Db     961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy    1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
Db    1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy    1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
Db    1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy    1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
Db    1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy    1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
Db    1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy    1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
Db    1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 3
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

```
  Query Match           100.0%;  Score 1311;  DB 5;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-269;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
```

```
Db        61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
Qy       121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
Db       121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
Qy       181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
Db       181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
Qy       241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
Db       241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
Qy       301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
Db       301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
Qy       361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
Db       361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
Qy       421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
Db       421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
Qy       481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
Db       481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
Qy       541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
Db       541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
Qy       601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
Db       601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
Qy       661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
Db       661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
Qy       721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
Db       721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
Qy       781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
Db       781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
Qy       841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
Db       841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
Qy       901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
Db       901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
Qy       961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
Db       961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
Qy      1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
Db      1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
Qy      1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
Db      1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
Qy      1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
Db      1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
Qy      1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
Db      1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
Qy      1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
Db      1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 4
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
XX
AC AAH76333;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 100.0%; Score 1311; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.2e-269;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
```

```
Qy       181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
Db       181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy       241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
Db       241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy       301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
Db       301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy       361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
Db       361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy       421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
Db       421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy       481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
Db       481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy       541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
Db       541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy       601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
Db       601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy       661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
Db       661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy       721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
Db       721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy       781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
Db       781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy       841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
Db       841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy       901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
Db       901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy       961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
Db       961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy      1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
Db      1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy      1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
Db      1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy      1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
Db      1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy      1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
Db      1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy      1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
Db      1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 5
AAH76340
ID AAH76340 standard; DNA; 255 BP.
XX
AC AAH76340;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 promoter fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; promoter; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Example 5; Fig 8; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a Z. mays Ms45 promoter fragment
XX
SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 12.3%; Score 160.6; DB 5; Length 255;
Best Local Similarity 97.6%; Pred. No. 1.5e-24;
Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1145 CGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGT 1204
Db         5 CCGCGGATCCCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGT 64

Qy      1205 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 1264
Db        65 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 124

Qy      1265 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
Db       125 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171
```

RESULT 6
AAH76334
ID AAH76334 standard; DNA; 158 BP.
XX
AC AAH76334;

```
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 5; Page 47; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a DNA fragment -38 to -195 bases upstream of
CC   the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC   nucleotide sequence
XX
SQ   Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

  Query Match             11.1%;  Score 146;  DB 5;  Length 158;
  Best Local Similarity   99.4%;  Pred. No. 1.7e-21;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 1213
             ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy       1214 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 1273
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120

Qy       1274 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
             |||||||||||||||||||||||||||||||||||||
Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158


RESULT 7
ADX48600
ID   ADX48600 standard; cDNA; 320 BP.
XX
AC   ADX48600;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 23340.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 23340; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This sequence represents a plant full length insert
CC   polynucleotide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 320 BP; 89 A; 53 C; 66 G; 112 T; 0 U; 0 Other;

  Query Match              5.6%;  Score 74;  DB 13;  Length 320;
  Best Local Similarity   73.1%;  Pred. No. 4.6e-06;
  Matches   95;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy        791 TTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAA 850
              |||| || |||||||||| || |   ||| ||||||   ||||| |||||   |||  ||
Db         62 TTTTTAGGGCTAGTTTGGAAATCTCATTTTTTTCAAGAGATTTTTATTTTCCTAAAGGAA 121

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
              |||||||||||||   || | |||||||||  | ||| | ||| ||||||| | | || |
Db        122 ATTAGTTTATTTTTCTTTGAGAAAATAGAATTCCCTTGGGAAATTAGAGTTTCTAAACGA 181

Qy        911 GCCCTAGAAT 920
              |||||  | |
Db        182 GCCCTTAATT 191
```

RESULT 8
ADX33444
ID ADX33444 standard; cDNA; 624 BP.
XX
AC ADX33444;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polynucleotide seqid 16264.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 16264; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 624 BP; 170 A; 119 C; 146 G; 189 T; 0 U; 0 Other;

Query Match 5.6%; Score 74; DB 13; Length 624;
Best Local Similarity 73.1%; Pred. No. 5.4e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

```
Qy      791 TTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAA 850
            |||| || |||||||||| || |  ||| ||||||   |||||  |||||   |||  ||
Db      374 TTTTTAGGGCTAGTTTGGAAATCTCATTTTTTTCAAGAGATTTTTATTTTCCTAAAGGAA 433

Qy      851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
            ||||||||||||    || | |||||||||  | ||| | ||| |||||||| | | ||| |
Db      434 ATTAGTTTATTTTTCTTTGAGAAAATAGAATTCCCTTGGGAAATTAGAGTTTCTAAACGA 493

Qy      911 GCCCTAGAAT 920
            |||||  | |
Db      494 GCCCTTAATT 503
```

RESULT 9
ADX61650
ID ADX61650 standard; cDNA; 2445 BP.
XX
AC ADX61650;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polynucleotide seqid 32493.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 32493; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of

```
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This sequence represents a plant full length insert
CC   polynucleotide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 2445 BP; 666 A; 470 C; 513 G; 796 T; 0 U; 0 Other;

  Query Match             5.6%;  Score 72.8;  DB 13;  Length 2445;
  Best Local Similarity   73.6%;  Pred. No. 1.4e-05;
  Matches  106;  Conservative    0;  Mismatches   37;  Indels    1;  Gaps    1;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTT-TCAAAGAATTTTGATTTTTTCAAAAAAA 850
              | | || |||||||||  |||||| ||| ||  ||||  ||||| ||||||  || ||||
Db       1623 TCTTAGGGCTAGTTTGAGAACCCTTTTTTTTCCCAAAAGATTTTCATTTTTCAAAGAAAA 1682

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
              ||||||| |||||  ||| | ||||||| || | |||| |||||||| |||  || ||||
Db       1683 ATTAGTTCATTTTTCCTTGAGAAAATAGGAATCCCTTAAAAAAATAGTGTTTTCAAACTA 1742

Qy        911 GCCCTAGAATGTTTTCCCAATAAA 934
              |||||  |   ||||    || ||
Db       1743 GCCCTTAATGTTTTTTTTCATGAA 1766


RESULT 10
ADX34996/c
ID   ADX34996 standard; cDNA; 928 BP.
XX
AC   ADX34996;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 17816.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 17816; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This sequence represents a plant full length insert
CC   polynucleotide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 928 BP; 245 A; 209 C; 219 G; 255 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 71.8;  DB 13;  Length 928;
  Best Local Similarity   78.0%;  Pred. No. 1.8e-05;
  Matches   99;  Conservative    0;  Mismatches   27;  Indels    1;  Gaps    1;

Qy        790 TTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
              |||||||||||| ||||||||| ||| || ||  || | ||||  |||||   || |||
Db        804 TTTTTAAGAGCTGGTTTGGCAA-CCTCATTTTTCTAAGGGATTTCTATTTTACTAAGAAA 746

Qy        850 AATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACT 909
              |||||||| ||||||||||   |||||| |||   ||| |||||||||||| ||| | |
Db        745 AATTAGTTAATTTTCTCTTAGAAAAATAAAAATGTCTTGGAAAAATAGAGTTACCAAATT 686

Qy        910 AGCCCTA 916
              | |||||
Db        685 ACCCCTA 679


RESULT 11
ADX61114
ID   ADX61114 standard; cDNA; 2537 BP.
XX
AC   ADX61114;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 31957.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
```

```
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 31957; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This sequence represents a plant full length insert
CC   polynucleotide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 2537 BP; 685 A; 484 C; 524 G; 844 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 71.8;  DB 13;  Length 2537;
  Best Local Similarity   73.4%;  Pred. No. 2.2e-05;
  Matches  105;  Conservative    0;  Mismatches   37;  Indels    1;  Gaps    1;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              | | || |||||||||  |||||| ||| | ||||  ||||| ||||||  || |||||
Db       1799 TCTTAGGGCTAGTTTGAGAACCCT-TTTTTCCCAAAAGATTTTCATTTTTCAAAGAAAAA 1857

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              |||||| |||||  ||| | ||||||| || | |||| |||||||| ||| || |||||
Db       1858 TTAGTTCATTTTTCCTTGAGAAAATAGGAATCCCTTAAAAAAATAGTGTTTTCAAACTAG 1917

Qy        912 CCCTAGAATGTTTTCCCAATAAA 934
              ||||  |  ||||    || ||
Db       1918 CCCTTAATGTTTTTTTTCATGAA 1940


RESULT 12
ADX60218
ID   ADX60218 standard; cDNA; 3607 BP.
XX
AC   ADX60218;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 31061.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 31061; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This sequence represents a plant full length insert
CC   polynucleotide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 3607 BP; 915 A; 695 C; 718 G; 1279 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 71.8;  DB 13;  Length 3607;
  Best Local Similarity   73.4%;  Pred. No. 2.4e-05;
  Matches  105;  Conservative    0;  Mismatches   37;  Indels    1;  Gaps    1;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              | | || |||||||||  |||||| ||| | ||||  ||||| ||||||  || |||||
Db       2871 TCTTAGGGCTAGTTTGAGAACCCT-TTTTTCCCAAAAGATTTTCATTTTTCAAAGAAAAA 2929

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              |||||| |||||  ||| | ||||||| || | |||| |||||||| ||| || |||||
Db       2930 TTAGTTCATTTTTCCTTGAGAAAATAGGAATCCCTTAAAAAAATAGTGTTTTCAAACTAG 2989

Qy        912 CCCTAGAATGTTTTCCCAATAAA 934
              ||||  |  ||||    || ||
Db       2990 CCCTTAATGTTTTTTTTCATGAA 3012


RESULT 13
AAL15210/c
ID   AAL15210 standard; cDNA; 883 BP.
XX
AC   AAL15210;
XX
DT   07-DEC-2001  (first entry)
XX
```

DE Human breast cancer expressed polynucleotide 7667.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS Claim 1; Page 1378; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 883 BP; 322 A; 18 C; 23 G; 396 T; 0 U; 124 Other;

Query Match 5.4%; Score 70.2; DB 4; Length 883;
Best Local Similarity 39.2%; Pred. No. 3.8e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

Qy 419 TGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCA 478
Db 873 TGTTNANNNANACCTANACCTTTAAANATNTNNNANTTNAANNAAAAAAATATTTAAANN 814

Qy 479 CTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAA 538
Db 813 TANAAATTTTTTNAAAAACTAATTNTAANNATANTNAATTTTTNTAAAAAAANNNAAAAA 754

Qy 539 AACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACA 598
Db 753 TTAAANNTTTTNTAANTTATTAAACCAAAATTTTTTTTTTAAAAAAATTTTTTAAAANTT 694

Qy 599 AATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTG 658
Db 693 AATAANTTNAATTTTTTTTAATNAAAAAAAAAAATTTNAAATTTTAANCAANTNTTTTTT 634

Qy 659 TTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATA 718
Db 633 TTTNTTTNNATAAAAAAANTTTTTAAATAAAAAANNANTTTTTTTTTATNAANATATTTAA 574

Qy 719 AAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAA 778
Db 573 AAAAAAANNTTTTTTTTTTTATANTTTATAAAAAATTTTATTTNTNTTTTTCTNAAAANA 514

Qy 779 TCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATT 838
Db 513 AAAAAAAATTTTTTTNAAANTNTTTTTAAAACCTTNAATTNAAANANAATTTTTTTNNNTT 454

Qy 839 TTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGA 898
Db 453 NNANATTAAAAAATTTATTTTTTTTANTTCTATAAATTAAAANAANANNAATAAAATAAT 394

Qy 899 GTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATT 958
Db 393 TTTNAAAAAAATAATTTTATANTTATNNAANANAAATTTAAAANTNAANANAAAAATTTTA 334

Qy 959 TGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCT 1018
Db 333 NTTTAAATAAATAAAATAANTTTAAACNTCATTTNTAATTTTATAATTAAAAAAAAAATT 274

Qy 1019 ATTT 1022
Db 273 NTTT 270

RESULT 14
ACN85231/c
ID ACN85231 standard; DNA; 960 BP.
XX
AC ACN85231;
XX
DT 02-DEC-2004 (first entry)
XX
DE Breast cancer related marker, seq id 6381.
XX
KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
OS Homo sapiens.
XX
PN US2003099974-A1.
XX
PD 29-MAY-2003.
XX
PF 18-JUL-2002; 2002US-00198846.
XX
PR 18-JUL-2001; 2001US-0306220P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2003-787014/74.
XX
PT Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
PS Disclosure; SEQ ID NO 6381; 36pp; English.
XX
CC The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
SQ Sequence 960 BP; 340 A; 33 C; 39 G; 421 T; 0 U; 127 Other;

Query Match 5.4%; Score 70.2; DB 11; Length 960;
Best Local Similarity 39.2%; Pred. No. 3.9e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

Qy 419 TGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCA 478

```
Db        933 TGTTNANNNANACCTANACCTTTAAANATNTNNNANTTNAANNAAAAAAATATTTAAANN 874

Qy        479 CTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAA 538

Db        873 TANAAATTTTTTNAAAAACTAATTNTAANNATANTNAATTTTTNTAAAAAAANNNAAAAA 814

Qy        539 AACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACA 598

Db        813 TTAAANNTTTTNTAANTTATTAAACCAAAATTTTTTTTTTAAAAAAATTTTTTAAAANTT 754

Qy        599 AATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTG 658

Db        753 AATAANTTNAATTTTTTTTAATNAAAAAAAAAAATTTNAAATTTTAANCAANTNTTTTTT 694

Qy        659 TTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATA 718

Db        693 TTTNTTTNNATAAAAAAANTTTTTAAATAAAAAANNANTTTTTTTTATNAANATATTTAA 634

Qy        719 AAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAA 778

Db        633 AAAAAAANNTTTTTTTTTTTATANTTTATAAAAAATTTTATTTNTNTTTTTCTNAAAANA 574

Qy        779 TCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATT 838

Db        573 AAAAAAAATTTTTTNAAANTNTTTTTAAAACCTTNAATTNAAANANAATTTTTTTNNNTT 514

Qy        839 TTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGA 898

Db        513 NNANATTAAAAAATTTATTTTTTTTTANTTCTATAAATTAAAANAANANNAATAAAATAAT 454

Qy        899 GTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATT 958

Db        453 TTTNAAAAAATAATTTTATANTTATNNAANANAAATTTAAAANTNAANANAAAAATTTTA 394

Qy        959 TGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCT 1018

Db        393 NTTTAAATAAATAAAATAANTTTAAACNTCATTTNTAATTTTATAATTAAAAAAAAAATT 334

Qy       1019 ATTT 1022

Db        333 NTTT 330
```

RESULT 15
ADR04296
ID ADR04296 standard; DNA; 13400 BP.
XX
AC ADR04296;
XX
DT 04-NOV-2004 (first entry)
XX
DE Corn FT homologue nucleotide sequence SEQ ID NO:63.
XX
KW flowering locus T gene; FT; terminal flower; TFL; Apetala3; Ap3; plant;
KW floral development; plant sterility; plant fertility; flowering time;
KW plant growth rate; inflorescence architecture; tissue culture morphology;
KW cell division; FT homologue; gene; ds.
XX
OS Zea mays.
XX
PN WO2004067723-A2.
XX
PD 12-AUG-2004.
XX
PF 29-JAN-2004; 2004WO-US002422.
XX
PR 30-JAN-2003; 2003US-00343477.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Danilevskaya O, Hermon P, Bruggemann E, Shirbroun D, Ananiev E;
PI Rafalski JA, Sakai H, Cahoon E, Cahoon R, Klein T;
XX
DR WPI; 2004-580996/56.
XX
PT New polynucleotides, specifically nucleic acid fragments encoding
PT flowering locus T gene (FT) or terminal flower (TFL), or Apetala3 (Ap3)
PT homologs, useful for floral development, e.g. engineering plant flowering
PT time.
XX
PS Claim 6; SEQ ID NO 63; 109pp; English.
XX
CC The present invention describes an isolated polynucleotide comprising a
CC first, second, third, fourth or fifth nucleotide sequence, or their
CC complement encoding a polypeptide either having flowering locus T gene
CC (FT), terminal flower (TFL), or Apetala3 (Ap3) homologue activity. Also
CC described: (1) a vector comprising the polynucleotide; (2) a recombinant
CC DNA construct comprising the polynucleotide; (3) transforming a cell by
CC transforming a cell with the polynucleotide; (4) a cell comprising the
CC recombinant DNA construct; (5) producing a plant comprising transforming
CC a plant cell with the polynucleotide, and regenerating a plant from the
CC transformed plant cell; (6) a plant comprising the recombinant DNA
CC construct; (7) a seed comprising the recombinant DNA construct; (8) an
CC isolated polynucleotide comprising a first nucleotide sequence, where the
CC first nucleotide sequence contains at least 30 nucleotides, and where the
CC first nucleotide sequence is comprised by another polynucleotide, where
CC the other polynucleotide includes the second, third, fourth, fifth or
CC sixth nucleotide sequence; (9) an isolated polypeptide having FT or Ap3
CC homologue activity, as described above; and (10) isolating a polypeptide
CC encoded by the polynucleotide comprising isolating the polypeptide from a
CC cell containing a recombinant DNA construct comprising the polynucleotide
CC operably linked to a regulatory sequence. The polynucleotides are useful
CC for floral development, e.g. engineering plant sterility/fertility,
CC flowering time, plant growth rate, inflorescence architecture, and tissue
CC culture morphology and the rate of cell division to enhance
CC transformation. The present sequence represents an FT homologue
CC nucleotide sequence from the present invention.
XX
SQ Sequence 13400 BP; 3962 A; 2948 C; 2695 G; 3795 T; 0 U; 0 Other;

Query Match 5.3%; Score 69.8; DB 13; Length 13400;
Best Local Similarity 73.6%; Pred. No. 9e-05;
Matches 89; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```
Qy        811 ACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTA 870

Db       8559 ACCATTTTATTTTCAAGAGGTTTTTATTTTATCAAGAAAAATTAGTTCATTTTCTCTTGG 8618

Qy        871 TAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAA 930

Db       8619 AAAAATAAAAATCCATTAGAAAAATGGGGTTGTCAAACTAGTCCTTATTTAGTTTTCCAT 8678

Qy        931 T 931

Db       8679 T 8679
```

Search completed: March 5, 2006, 18:11:41
Job time : 873.437 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 18:11:59 ; Search time 6189.32 Seconds
(without alignments)
9910.279 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311
Perfect score: 1311
Sequence: 1 ccatggtgtctctatgaaaa.....cgcacatatgtttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	908	69.3	963	9	CC656933	CC656933 OGWDQ20TM
	2	679	51.8	915	10	CG224225	CG224225 OG1AG08TV
c	3	419.8	32.0	687	9	CC656939	CC656939 OGWDQ20TV
c	4	96	7.3	715	10	CG252571	CG252571 OG4BB05TC
	5	92	7.0	967	10	CL235046	CL235046 ZMMBBb057
	6	91	6.9	860	10	CZ295176	CZ295176 ZMMBF0063
	7	88	6.7	814	10	CG048704	CG048704 PUILO19TB
c	8	87.8	6.7	754	10	CG414922	CG414922 ZMMBBb029
c	9	86.8	6.6	950	9	CC439901	CC439901 PUHRV15TB
	10	86.6	6.6	652	9	CC384247	CC384247 PUHOC67TB
c	11	86.6	6.6	797	9	CC400575	CC400575 PUHLU61TD
	12	86.6	6.6	820	9	CC400574	CC400574 PUHLU61TB
	13	85.4	6.5	471	10	CG103452	CG103452 PUJBB19TB
c	14	85.2	6.5	765	10	CG082135	CG082135 PUFQX12TD
c	15	85.2	6.5	781	9	CC630219	CC630219 OGUCG53TV
c	16	85.2	6.5	815	10	CG349565	CG349565 OG0FI83TH
	17	85.2	6.5	834	9	CC630210	CC630210 OGUCG53TH
	18	85	6.5	789	9	CC433618	CC433618 PUHHP17TD
c	19	84.6	6.5	1092	10	CNS020K7	AL175696 Tetraodon
	20	84	6.4	793	9	BZ816381	BZ816381 PUFBA86TD
c	21	84	6.4	1078	10	CL997678	CL997678 ZMMBHf001
	22	83.8	6.4	530	10	CG201774	CG201774 PUICH24TB
c	23	83.6	6.4	722	10	CG333914	CG333914 OG0AD14TH
	24	83.6	6.4	722	10	CG333929	CG333929 OG0AD14TV
	25	83.6	6.4	781	10	CG034985	CG034985 PUIGR68TB
	26	83.6	6.4	861	9	BZ797976	BZ797976 PUGBP28TB
c	27	83.6	6.4	865	9	CC430754	CC430754 PUHEP05TB
c	28	81.8	6.2	861	10	CG102092	CG102092 PUFYW94TD
c	29	81.8	6.2	947	9	CC435780	CC435780 PUHSN02TD
	30	81.8	6.2	981	9	BZ784278	BZ784278 PUFHV09TD
c	31	81.8	6.2	1017	9	CC620594	CC620594 OGUCC26TV
	32	81.6	6.2	649	9	CC613918	CC613918 OGJAG88TV
c	33	81.6	6.2	733	9	BZ778636	BZ778636 ih02f10.g
c	34	81.6	6.2	925	10	CG071791	CG071791 PUJBK02TB
c	35	81.6	6.2	960	9	BZ676889	BZ676889 PUBIG17TD
	36	81.6	6.2	994	9	CC003943	CC003943 PUDJN65TD
	37	81.6	6.2	1016	10	CL996481	CL996481 ZMMBHf000
	38	81.6	6.2	1022	10	CL984151	CL984151 ZMMBHd000
	39	81.4	6.2	999	9	CC385762	CC385762 PUHFK14TD
	40	81.2	6.2	756	10	CG221693	CG221693 OGWMG04TH
c	41	81	6.2	1101	10	CNS00LT2	AL078714 Drosophil
c	42	80.8	6.2	792	8	DR794756	DR794756 ZM_BFb001
	43	80.6	6.1	699	9	BZ996930	BZ996930 PUGIR50TB
c	44	80.6	6.1	832	9	CC676387	CC676387 OGKAS81TH
c	45	80.6	6.1	839	10	CG246849	CG246849 OGWAL22TH

ALIGNMENTS

RESULT 1
CC656933
LOCUS CC656933 963 bp DNA linear GSS 19-JUN-2003
DEFINITION OGWDQ20TM ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15, genomic survey sequence.
ACCESSION CC656933
VERSION CC656933.1 GI:32060225
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 963)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGWDQ20TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: methylation filtered.
FEATURES Location/Qualifiers
source 1. .963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0554D15"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 69.3%; Score 908; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 9.6e-180;
Matches 908; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 404 ACCAGCCATGTCTAGTGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACC 463

```
Db          1 ACCAGCCATGTCTAGTGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACC 60

Qy        464 AAAGTAACTAATCCACTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACC 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGTAACTAATCCACTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACC 120

Qy        524 ACGATCCTCTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAA 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACGATCCTCTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAA 180

Qy        584 CATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTC 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTC 240

Qy        644 ATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTT 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTT 300

Qy        704 TCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAG 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAG 360

Qy        764 ATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTT 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTT 420

Qy        824 CAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAAC 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAAC 480

Qy        884 ACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCA 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCA 540

Qy        944 CTGTGTATAATTATTTGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGA 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGTGTATAATTATTTGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGA 600

Qy       1004 AACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAG 660

Qy       1064 TTTTAAGTCATTGTCCCTGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCT 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTTAAGTCATTGTCCCTGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCT 720

Qy       1124 CATAGGTGCCTACGTCAGGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGC 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CATAGGTGCCTACGTCAGGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGC 780

Qy       1184 TTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGAT 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGAT 840

Qy       1244 ACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTT 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTT 900

Qy       1304 TCCTGAAC 1311
              ||||||||
Db        901 TCCTGAAC 908
```

```
RESULT 2
CG224225
LOCUS       CG224225                915 bp    DNA     linear   GSS 22-AUG-2003
DEFINITION  OG1AG08TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0716B15,
            genomic survey sequence.
ACCESSION   CG224225
VERSION     CG224225.1  GI:34124113
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1AG08TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0716B15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN
```

```
  Query Match             51.8%;  Score 679;  DB 10;  Length 915;
  Best Local Similarity   99.9%;  Pred. No. 9.5e-132;
  Matches  690;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        621 TGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTT 680
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db          1 TGACGTTCAGATTTTTCTTTTTC-TTCTTGTTATTTTGTTATTGTTTTTATATACATTTT 59

Qy        681 CTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTT 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 CTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTT 119

Qy        741 ATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGC 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 ATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGC 179

Qy        801 TAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTAT 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTAT 239

Qy        861 TTTCTCTTTATAAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAAT 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TTTCTCTTTATAAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAAT 299

Qy        921 GTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTGGCCAGCCCCATAAATTATTT 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTGGCCAGCCCCATAAATTATTT 359

Qy        981 AAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAG 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 AAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAG 419

Qy       1041 GGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCCTGAGATGTGCGGTTTGGCAA 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 GGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCCTGAGATGTGCGGTTTGGCAA 479

Qy       1101 CGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCAGGTTCGGCAGCTCTCGTGTC 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCAGGTTCGGCAGCTCTCGTGTC 539

Qy       1161 ATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCT 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        540 ATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCT 599

Qy       1221 TGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTT 1280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTT 659

Qy       1281 CCATGCAAACACGCACATATGTTTCCTGAAC 1311
             |||||||||||||||||||||||||||||||
Db        660 CCATGCAAACACGCACATATGTTTCCTGAAC 690


RESULT 3
CC656939/c
LOCUS       CC656939                687 bp    DNA     linear   GSS 19-JUN-2003
DEFINITION  OGWDQ20TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15,
            genomic survey sequence.
ACCESSION   CC656939
VERSION     CC656939.1  GI:32060231
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 687)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OGWDQ20TM
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .687
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0554D15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match           32.0%;  Score 419.8;  DB 9;  Length 687;
  Best Local Similarity   99.5%;  Pred. No. 2e-77;
  Matches  421;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        889 GAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTG 948
              |||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        687 GAAAAATAGAGTGCCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTG 628

Qy        949 TATAATTATTTGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCA 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TATAATTATTTGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCA 568

Qy       1009 AATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGAGAGAGAGGAAGAAATCAGTTTTA 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 AATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGAGAGAGAGGAAGAAATCAGTTTTA 508

Qy       1069 AGTCATTGTCCCTGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAG 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 AGTCATTGTCCCTGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAG 448

Qy       1129 GTGCCTACGTCAGGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTT 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        447 GTGCCTACGTCAGGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTT 388

Qy       1189 CAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTA 1248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTA 328

Qy       1249 CTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTG 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTG 268

Qy       1309 AAC 1311
              |||
Db        267 AAC 265


RESULT 4
CG252571/c
LOCUS       CG252571                715 bp    DNA     linear   GSS 25-AUG-2003
DEFINITION  OG4BB05TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0809B10,
            genomic survey sequence.
ACCESSION   CG252571
VERSION     CG252571.1  GI:34154661
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 715)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .715
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0809B10"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match           7.3%;  Score 96;  DB 10;  Length 715;
  Best Local Similarity   80.1%;  Pred. No. 1.5e-09;
  Matches  125;  Conservative    0;  Mismatches   30;  Indels    1;  Gaps    1;

Qy        135 TCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCATCTGACCGAGGCCCA 194
              || | ||||     ||| | ||||||||||||||| ||||||| ||||| |||||||||
Db        356 TCACACACCCGCTCTTGTGCCTTTCTTGGATGTCGGCGGTTCCCGTCTGATCGAGGCCCA 297

Qy        195 TCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCACGAGACGTATCGGG 254
              | || ||||| ||||   |||||| || |||||||||||||||||| ||| |   |||
Db        296 TAAGGCACCTGTCGGAGAACCCATTAAAGGCCTTTCGGATGGCCCACGGGACAT-CGGGG 238

Qy        255 TCGTGGTGATCCAGGGGATATATGTCCCCCACAATC 290
              ||||||||| || ||||||||| | |||||||||| |
Db        237 TCGTGGTGACCCGGGGGATATCTATCCCCCACAACC 202


RESULT 5
CL235046
```

```
LOCUS       CL235046                967 bp    DNA     linear   GSS 15-JAN-2004
DEFINITION  ZMMBBb0575O01r ZMMBBb (HindIII) Zea mays genomic clone
            ZMMBBb0575O01 3', genomic survey sequence.
ACCESSION   CL235046
VERSION     CL235046.1  GI:40891729
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 967)
  AUTHORS   Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
            Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
  TITLE     Sequencing of the maize genome at PGIR (2003c)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Bharti,A.K.
            Dr.Joachim Messing's lab
            The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
            University
            190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            Tel: 732 445 3801
            Fax: 732 445 5735
            Email: bharti@waksman.rutgers.edu
            Seq primer: SP6
            Class: BAC ends
            High quality sequence start: 64.
FEATURES             Location/Qualifiers
     source          1. .967
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0575O01"
                     /lab_host="E. coli DH10B"
                     /clone_lib="ZMMBBb (HindIII)"
                     /note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             7.0%;  Score 92;  DB 10;  Length 967;
  Best Local Similarity   78.6%;  Pred. No. 1e-08;
  Matches  110;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy        150 TGCACATTTCTTGGATGTCGGTGGTTCCCATCTGACCGAGGCCCATCAGACACCTTTCGG 209
              ||| | ||||||||||  ||| |||||||| |||||  ||||||||||    ||||||| ||
Db        827 TGCGCCTTTCTTGGATACCGGCGGTTCCCGTCTGATTGAGGCCCATCCAGCACCTTTTGG 886

Qy        210 GACACCCATCAAGGGCCTTTCGGATGGCCCACGAGACGTATCGGGTCGTGGTGATCCAGG 269
              |   |||||||| |||||||||||| ||||| || ||||| || |   |||| |   ||
Db        887 GTGGCCCATCAAGTGCCTTTCGGATGGTCCACGGGATGTATCCGGGCCGTGTGACCTGGG 946

Qy        270 GGATATATGTCCCCCACAAT 289
              |||||| |||||||||| | |
Db        947 GGATATCTGTCCCCCATAGT 966


RESULT 6
CZ295176
LOCUS       CZ295176                860 bp    DNA     linear   GSS 23-MAR-2005
DEFINITION  ZMMBF0063M20f ZMMBF Zea mays genomic clone ZMMBF0063M20 5', genomic
            survey sequence.
ACCESSION   CZ295176
VERSION     CZ295176.1  GI:61708952
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 860)
  AUTHORS   Bharti,A.K., Nelson,A.B., Young,S., Keizer,G., Zohovetz,V., Fuks,G.
            and Messing,J.
  TITLE     Construction, Sequencing and Characterization of a Fosmid Library
            of the B73 Maize Genome
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Bharti,A.K.
            Dr.Joachim Messing's lab
            The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
            University
            190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            Tel: 732 445 3801
            Fax: 732 445 5735
            Email: bharti@waksman.rutgers.edu
            Seq primer: Fos_F
            Class: fosmid ends.
FEATURES             Location/Qualifiers
     source          1. .860
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBF0063M20"
                     /lab_host="EPI100-T1"
                     /clone_lib="ZMMBF"
                     /note="Vector: pEpiFOS-5; Site_1: Eco72I"
ORIGIN

  Query Match             6.9%;  Score 91;  DB 10;  Length 860;
  Best Local Similarity   62.6%;  Pred. No. 1.7e-08;
  Matches  142;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy        782 TTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTT 841
              |||  |    | |||| ||||||||| ||||||| ||| ||  || ||||||| ||||||
Db         51 TTCAAAGCACTCTAAGGGCTAGTTTGACAACCCTATTTTTTCTAAGGAATTTTTATTTTT 110

Qy        842 TCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTT 901
              | || |||||||||||| |||||||||    ||||| |||| | |||| ||||||| ||||
Db        111 TAAAGAAAAATTAGTTCATTTTCTCTTGTGAAAATGGAAATCCCTTAAAAAAATAAAGTT 170

Qy        902 GCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTGG 961
               | | | ||||| || ||  | |   |  |    | ||| |  ||| || |  |   |
Db        171 ACTAAATTAGCCTTAAAAGTTGGTTTGATGACCACAAAATTAGAGTGGATTAAGAGACGA 230

Qy        962 CCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCA 1008
                | |||  |       |  |  |  |  ||| |||     | |||
Db        231 GAAATCCCTTGCTAATGATTGAGGGGGAGAGAATTCGTAACTATCCA 277


RESULT 7
CG048704
LOCUS       CG048704                814 bp    DNA     linear   GSS 19-AUG-2003
DEFINITION  PUILO19TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0611C13,
            genomic survey sequence.
ACCESSION   CG048704
VERSION     CG048704.1  GI:33920884
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 814)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUILO19TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
```

```
FEATURES             Location/Qualifiers
     source          1. .814
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa0611C13"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match           6.7%;  Score 88;  DB 10;  Length 814;
  Best Local Similarity  62.7%;  Pred. No. 7.1e-08;
  Matches  153;  Conservative    0;  Mismatches   90;  Indels    1;  Gaps    1;

Qy        713 TTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTT 772
              |||| |  | |    || ||       |||||     |   ||  ||||   | |
Db        487 TTTAGAGGACGTTGGTAGAGATGGAAAGATATAGAGAAAATAATCTTTTAGAGAATGTTA 546

Qy        773 CAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATT 832
              |||   |       ||| ||||||||||||||| || ||  ||| ||||||| |||
Db        547 TAAAGGACAAGAAACATTTCTTAAGAGCTAGTTTGGTAATCCCATTTTTTTCAAAGGATT 606

Qy        833 TTGATTTTTTCAAAA-AAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAA 891
              || ||| |  |||   |||| |||||||||| |  ||| |||||| | || ||||  |||
Db        607 TTCATTCTCACAAGGGAAAAATAGTTTATTTCCCTTTTGTAAAATGGTAATCACTCGGAA 666

Qy        892 AAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTAT 951
              ||||||||||  || ||||||||||   |  | |||| ||||     | ||| |  | | |
Db        667 AAATAGAGTTCTCAAACTAGCCCTAAGGTTATATCCCGATAATCCCCTATCTCATTCTCT 726

Qy        952 AATT 955
              | ||
Db        727 ATTT 730


RESULT 8
CG414922/c
LOCUS       CG414922                754 bp    DNA     linear   GSS 08-SEP-2003
DEFINITION  ZMMBBb0290L09.r ZMMBBb Zea mays genomic clone ZMMBBb0290L09 3',
            genomic survey sequence.
ACCESSION   CG414922
VERSION     CG414922.1  GI:34505144
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 754)
  AUTHORS   Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
            and Wing,R.
  TITLE     Sequencing of the maize genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: T7
            BACKWARD: M13r
            Plate: 0290  row: L  column: 09
            Seq primer: M13r
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .754
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0290L09"
                     /lab_host="DH10B"
                     /clone_lib="ZMMBBb"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII; Zea mays L. ssp. mays"
ORIGIN

  Query Match           6.7%;  Score 87.8;  DB 10;  Length 754;
  Best Local Similarity  71.2%;  Pred. No. 7.8e-08;
  Matches  116;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy        779 TCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATT 838
              |||    |  | ||  ||| |||||||||  || |   ||| |||| ||| ||||| |||
Db        680 TCTAAATTTTATTCCTAAAGGCTAGTTTGAAAAGCACATTTTTTTCCAAGGATTTTAATT 621

Qy        839 TTTTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGA 898
              |||  ||| |||||||  || |||||  |||| | ||||||||||  ||| ||||||||| |
Db        620 TTTCTAAATAAAATTAATTCATTTTTTCTTGAAAAAATAGAAATCACATAGAAAAATAAA 561

Qy        899 GTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAAT 941
              ||| | | |||| | ||| ||  | |||    | | |||||
Db        560 GTTCCTAAACTAACTCTAAAAAATATTCTTCGTCATTTACAGT 518


RESULT 9
CC439901/c
LOCUS       CC439901                950 bp    DNA     linear   GSS 20-MAY-2003
DEFINITION  PUHRV15TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa519D06,
            genomic survey sequence.
ACCESSION   CC439901
VERSION     CC439901.1  GI:30940477
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 950)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUHRV15TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .950
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa519D06"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match           6.6%;  Score 86.8;  DB 9;  Length 950;
  Best Local Similarity  72.4%;  Pred. No. 1.3e-07;
  Matches  126;  Conservative    0;  Mismatches   47;  Indels    1;  Gaps    1;

Qy        787 ATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAA 846
              || ||| |||   |||||||| |||| | ||| | |||  | ||||| ||||||||||
Db        241 ATCTTTCTAATGACTAGTTTGGGAACCTTATTT-TCTCACGGGATTTTTATTTTTTCAAT 183
```

```
Qy        847 AAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAG 906
              ||||||||||||||||||||| | |||||| ||| |||||   |||||| |||| |||
Db        182 GAAAATTAGTTTATTTTCTCTTGAGAAAATATAAATCACTTGAGAAAATATAGTTTCCAA 123

Qy        907 ACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              |||||||||| | | ||||   ||  |   | ||      || |||  ||  |||
Db        122 ACTAGCCCTAAAGTATTTTATAAAATAGAGAAAAAATTCGTCTATTTTTCATTG 69


RESULT 10
CC384247
LOCUS       CC384247                652 bp    DNA     linear   GSS 19-MAY-2003
DEFINITION  PUHOC67TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa495L13,
            genomic survey sequence.
ACCESSION   CC384247
VERSION     CC384247.1  GI:30864143
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 652)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUHOC67TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa495L13"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             6.6%;  Score 86.6;  DB 9;  Length 652;
  Best Local Similarity   63.9%;  Pred. No. 1.4e-07;
  Matches  131;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy        714 TTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTC 773
              | || || ||        ||  |  ||  ||  ||       |||| ||  |||   | |
Db        341 TAATTAATTGTAGTACTAGCTACATTCTTAATCGACATATTATGTAGTTAATTAATGCTA 400

Qy        774 AAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTT 833
              | ||   | | |  |  | ||||||||||||||||||| |||  | ||| | | | ||| |||
Db        401 AGAATCTTATTTATTAATGTTAAGAGCTAGTTTGGGAACTATATTTTTCTAAGAGATTTT 460

Qy        834 TGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAA 893
              |  |||||||||   |||||||||||||||  | |   ||||||||||  | || |||||
Db        461 TATTTTTTTCAAGTGAAATTAGTTTATTTTTCCATGGGAAAATAGAAATTAATTGGAAAA 520

Qy        894 ATAGAGTTGCCAGACTAGCCCTAGA 918
              ||   ||| ||| ||||||||||| |
Db        521 ATTAGGTTTCCAAACTAGCCCTAAA 545


RESULT 11
CC400575/c
LOCUS       CC400575                797 bp    DNA     linear   GSS 19-MAY-2003
DEFINITION  PUHLU61TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa480L01,
            genomic survey sequence.
ACCESSION   CC400575
VERSION     CC400575.1  GI:30880665
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 797)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUHLU61TB
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .797
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa480L01"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             6.6%;  Score 86.6;  DB 9;  Length 797;
  Best Local Similarity   59.7%;  Pred. No. 1.4e-07;
  Matches  163;  Conservative    0;  Mismatches  109;  Indels    1;  Gaps    1;

Qy        713 TTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTT 772
              |||| |  | |   || ||        |||||     |   || | ||||  | |
Db        577 TTTAGAGGACGTTGGTAGAGATGGAAAGATATAGAGAAAATAATCTTTTAGAGAATGTTA 518

Qy        773 CAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATT 832
              |||   |        |||||||||||||||||||| || ||  ||| |||||||| |||
Db        517 TAAAGGACAAGAAACATTTTTTAAGAGCTAGTTTGGTAATCCCATTTTTTTCAAAGGATT 458

Qy        833 TTGATTTTTTCAAAA-AAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAA 891
              || ||| |  |||   |||| ||||||||| |  ||| |||||| | || ||||  |||
Db        457 TTCATTCTCACAAGGGAAAAATAGTTTATTTCCCTTTTGTAAAATGGTAATCACTCGGAA 398

Qy        892 AAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTAT 951
              ||||||||||  || ||||||||||   | | ||| | ||    | ||| |  | | |
Db        397 AAATAGAGTTCTCAAACTAGCCCTAAGGTTATATCCGATAAATCCCCTATCTCATTCTCT 338

Qy        952 AATTATTTGGCCAGCCCCATAAATTATTTAAAC 984
              | ||     |||       ||||  | | | |
Db        337 ATTTCAAATTCCACTTTGCAAAATAGTGTCATC 305


RESULT 12
CC400574
LOCUS       CC400574                820 bp    DNA     linear   GSS 19-MAY-2003
DEFINITION  PUHLU61TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa480L01,
            genomic survey sequence.
ACCESSION   CC400574
VERSION     CC400574.1  GI:30880664
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
```

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 820)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUHLU61TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .820
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa480L01"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             6.6%;  Score 86.6;  DB 9;  Length 820;
  Best Local Similarity   59.7%;  Pred. No. 1.4e-07;
  Matches  163;  Conservative    0;  Mismatches  109;  Indels    1;  Gaps    1;

Qy        713 TTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTT 772
              |||| |  | |    || ||        |||||      |   || | ||||  |  |
Db        374 TTTAGAGGACGTTGGTAGAGATGGAAAGATATAGAGAAAATAATCTTTTAGAGAATGTTA 433

Qy        773 CAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATT 832
              |||   |       ||||||||||||||||||||| || ||  ||| |||||||| |||
Db        434 TAAAGGACAAGAAACATTTTTTAAGAGCTAGTTTGGTAATCCCATTTTTTTCAAAGGATT 493

Qy        833 TTGATTTTTTCAAAA-AAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAA 891
              || ||| |  |||   |||| ||||||||| |  ||| |||||| | || ||||  |||
Db        494 TTCATTCTCACAAGGGAAAAATAGTTTATTTCCCTTTTGTAAAATGGTAATCACTCGGAA 553

Qy        892 AAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTAT 951
              ||||||||||  || ||||||||||   |  | ||| |  ||    | ||| |  | | |
Db        554 AAATAGAGTTCTCAAACTAGCCCTAAGGTTATATCCGATAAATCCCCTATCTCATTCTCT 613

Qy        952 AATTATTTGGCCAGCCCCATAAATTATTTAAAC 984
              | ||      |||     ||||  | | | |
Db        614 ATTTCAAATTCCACTTTGCAAAATAGTGTCATC 646


RESULT 13
CG103452
LOCUS       CG103452               471 bp    DNA     linear   GSS 20-AUG-2003
DEFINITION  PUJBE19TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0628D13,
            genomic survey sequence.
ACCESSION   CG103452
VERSION     CG103452.1  GI:33985746
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 471)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUJBE19TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .471
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa0628D13"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             6.5%;  Score 85.4;  DB 10;  Length 471;
  Best Local Similarity   69.5%;  Pred. No. 2.4e-07;
  Matches  116;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy        789 TTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAA 848
              |||| | |||||||||||||| || ||  ||| || ||  || ||    |||||||||
Db        216 TTTTCTTAGAGCTAGTTTGGAAATCCCATTTTTTCCAGGGATTTACATTTTTTTCAAGGG 275

Qy        849 AAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGAC 908
              ||||||||| ||||||  || | |||||||| || ||| | |  |||||| ||| ||| ||
Db        276 AAATTAGTTCATTTTCCTTTGAAAAAATAGGAATCACATGGGGAAATAGTGTTCCCAAAC 335

Qy        909 TAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATT 955
              |||| ||.  |   ||| | || ||| || |  || ||  ||  |||
Db        336 TAGCGCTTAGACTATTTTCTAAAAAAGTAGATCCAATGAATAAGATT 382


RESULT 14
CG082135/c
LOCUS       CG082135               765 bp    DNA     linear   GSS 20-AUG-2003
DEFINITION  PUFQX12TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0712B24,
            genomic survey sequence.
ACCESSION   CG082135
VERSION     CG082135.1  GI:33964429
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUFQX12TB
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
```

```
                     /clone="ZMMBTa0712B24"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             6.5%;  Score 85.2;  DB 10;  Length 765;
  Best Local Similarity   71.8%;  Pred. No. 2.7e-07;
  Matches  125;  Conservative    0;  Mismatches   48;  Indels    1;  Gaps    1;

Qy        787 ATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAA 846
              || ||| |||   |||||||||| |||| | ||| | |||  | ||||| ||||||||||
Db        255 ATCTTTCTAATGACTAGTTTGGGAACCTTATTT-TCTCACGGGATTTTTATTTTTTCAAT 197

Qy        847 AAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAG 906
              |||||||||||||||||||| | |||||| ||| |||||   |||||| |||| |||
Db        196 GAAAATTAGTTTATTTTCTCTTGAGAAAATATAAATCACTTGAGAAAATATAGTTTCCAA 137

Qy        907 ACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              |||||||||| | | ||||   ||  |     ||     || |||  ||  |||
Db        136 ACTAGCCCTAAAGTATTTTATAAAATAGAGGAAAAATTCGTCTATTTTTCATTG 83


RESULT 15
CC630219/c
LOCUS       CC630219                781 bp    DNA     linear   GSS 19-JUN-2003
DEFINITION  OGUCG53TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0406J09,
            genomic survey sequence.
ACCESSION   CC630219
VERSION     CC630219.1  GI:32003354
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 781)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OGUCG53TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .781
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0406J09"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             6.5%;  Score 85.2;  DB 9;  Length 781;
  Best Local Similarity   71.8%;  Pred. No. 2.7e-07;
  Matches  125;  Conservative    0;  Mismatches   48;  Indels    1;  Gaps    1;

Qy        787 ATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAA 846
              || ||| |||   |||||||||| |||| | ||| | |||  | ||||| ||||||||||
Db        665 ATCTTTCTAATGACTAGTTTGGGAACCTTATTT-TCTCACGGGATTTTTATTTTTTCAAT 607

Qy        847 AAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAG 906
              |||||||||||||||||||| | |||||| ||| |||||   |||||| |||| |||

Db        606 GAAAATTAGTTTATTTTCTCTTGAGAAAATATAAATCACTTGAGAAAATATAGTTTCCAA 547

Qy        907 ACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              |||||||||| | | ||||   ||  |     ||     || |||  ||  |||
Db        546 ACTAGCCCTAAAGTATTTTATAAAATAGAGGAAAAATTCGTCTATTTTTCATTG 493


Search completed: March  6, 2006, 01:57:48
Job time : 6189.32 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 21:54:12 ; Search time 366.534 Seconds
(without alignments)
6357.883 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311
Perfect score: 1311
Sequence: 1 ccatggtgtctctatgaaaa.....cgcacatatgtttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1311	100.0	1394	3	US-08-880-499-1	Sequence 1, Appli
	2	1311	100.0	1394	3	US-08-880-499-2	Sequence 2, Appli
c	3	66.4	5.1	6027	2	US-08-968-542C-1	Sequence 1, Appli
c	4	66.4	5.1	6027	3	US-09-554-467A-1	Sequence 1, Appli
c	5	58.6	4.5	2523	2	US-08-410-784A-3	Sequence 3, Appli
c	6	57	4.3	19124	2	US-08-487-826B-13	Sequence 13, Appl
c	7	56.4	4.3	50000	3	US-09-662-254B-26	Sequence 26, Appl
	8	55.8	4.3	279	3	US-09-313-294A-5397	Sequence 5397, Ap
c	9	55.6	4.2	612	3	US-09-902-540-1357	Sequence 1357, Ap
c	10	55.6	4.2	1394	3	US-08-880-499-1	Sequence 1, Appli
c	11	55.6	4.2	1394	3	US-08-880-499-2	Sequence 2, Appli
	12	55.4	4.2	1055	3	US-09-806-708B-23	Sequence 23, Appl
c	13	55.2	4.2	2614	3	US-09-004-056-1	Sequence 1, Appli
	14	55	4.2	396	3	US-09-640-173-53	Sequence 53, Appl
	15	55	4.2	396	3	US-09-713-550-53	Sequence 53, Appl
	16	55	4.2	396	3	US-09-825-294-53	Sequence 53, Appl
	17	55	4.2	396	3	US-09-970-966-53	Sequence 53, Appl
c	18	54.6	4.2	55886	3	US-09-949-016-15129	Sequence 15129, A
	19	53.8	4.1	307	3	US-09-313-294A-4743	Sequence 4743, Ap
	20	53.8	4.1	6027	2	US-08-968-542C-1	Sequence 1, Appli
	21	53.8	4.1	6027	3	US-09-554-467A-1	Sequence 1, Appli
	22	52.2	4.0	1141	3	US-09-806-708B-22	Sequence 22, Appl
c	23	51.6	3.9	134987	3	US-09-949-016-15348	Sequence 15348, A
c	24	51.6	3.9	134987	3	US-09-949-016-15349	Sequence 15349, A
c	25	51.6	3.9	134987	3	US-09-949-016-15350	Sequence 15350, A
c	26	51.6	3.9	134987	3	US-09-949-016-15507	Sequence 15507, A
c	27	51.6	3.9	134987	3	US-09-949-016-15508	Sequence 15508, A
c	28	51.6	3.9	134987	3	US-09-949-016-15509	Sequence 15509, A
c	29	51.4	3.9	50383	3	US-09-949-016-17600	Sequence 17600, A
	30	51.4	3.9	129415	3	US-09-949-016-16997	Sequence 16997, A
c	31	51	3.9	601	3	US-09-949-016-156535	Sequence 156535,
c	32	51	3.9	1141	3	US-09-806-708B-22	Sequence 22, Appl
c	33	50.6	3.9	1039	3	US-09-902-540-1280	Sequence 1280, Ap
	34	50.6	3.9	16573	3	US-09-949-016-14876	Sequence 14876, A
c	35	50.6	3.9	18773	3	US-09-949-016-14164	Sequence 14164, A
	36	50.4	3.8	731	2	US-08-451-405A-2	Sequence 2, Appli
c	37	50.2	3.8	832	3	US-09-621-976-2813	Sequence 2813, Ap
c	38	50.2	3.8	317366	3	US-09-949-016-16001	Sequence 16001, A
	39	50	3.8	1392	3	US-09-257-584-1	Sequence 1, Appli
c	40	49.8	3.8	2435	3	US-09-306-593-1	Sequence 1, Appli
	41	49.8	3.8	231129	3	US-09-949-016-16110	Sequence 16110, A
	42	49.8	3.8	266293	3	US-09-949-016-11934	Sequence 11934, A
	43	49.6	3.8	50000	3	US-09-662-254B-23	Sequence 23, Appl
c	44	49.6	3.8	640681	3	US-09-790-988-1	Sequence 1, Appli
c	45	49.2	3.8	126176	3	US-09-949-016-16137	Sequence 16137, A

ALIGNMENTS

RESULT 1
US-08-880-499-1
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-880-499-1

Query Match 100.0%; Score 1311; DB 3; Length 1394;

```
  Best Local Similarity   100.0%;  Pred. No. 8.7e-305;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 2
US-08-880-499-2
; Sequence 2, Application US/08880499
; Patent No. 6037523
;  GENERAL INFORMATION:
;    APPLICANT:  Albertson, Marc C.
;    APPLICANT:  Fox, Tim W.
;    APPLICANT:  Carl, Garnaat W.
;    APPLICANT:  Huffman, Gary A.
;    APPLICANT:  Kendall, Timmy L.
;    TITLE OF INVENTION:  MALE TISSUE-PREFERRED REGULATORY REGION
;    TITLE OF INVENTION:  AND METHOD OF USING SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PIONEER HI-BRED INTERNATIONAL, INC.
;      STREET:  Darwin Building, 7100 N.W. 62nd Ave., P.O.
;      STREET:  Box 1000
;      CITY:  Johnston
;      STATE:  Iowa
;      COUNTRY:  USA
;      ZIP:  50131
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/880,499
;      FILING DATE:  CONCURRENTLY HEREWITH
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sweeney, Patricia A.
;      REGISTRATION NUMBER:  32,733
;      REFERENCE/DOCKET NUMBER:  0578
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 248-4800
;      TELEFAX:  (515) 248-4844
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1394 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-880-499-2

  Query Match             100.0%;  Score 1311;  DB 3;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-305;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
```

```
Qy        61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
Db        61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy       121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
Db       121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy       181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
Db       181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy       241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
Db       241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy       301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
Db       301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy       361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
Db       361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy       421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
Db       421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy       481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
Db       481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy       541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
Db       541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy       601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
Db       601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy       661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
Db       661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy       721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
Db       721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy       781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
Db       781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy       841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
Db       841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy       901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
Db       901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy       961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
Db       961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy      1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
Db      1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy      1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
Db      1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy      1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
Db      1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy      1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
Db      1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy      1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
Db      1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 3
US-08-968-542C-1/c
; Sequence 1, Application US/08968542C
; Patent No. 5981728
; GENERAL INFORMATION:
; APPLICANT: Myers, et al.
; TITLE OF INVENTION: dull1 Codes For A No. 5981728el Starch
; TITLE OF INVENTION: Synthase
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 6.0.1 for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,542C
; FILING DATE: No. 5981728ember 12, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D6036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6027 bp
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to mRNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: maize
; TISSUE TYPE: endosperm
; IMMEDIATE SOURCE:
; LIBRARY: maize endosperm cDNA library in
; LIBRARY: (gt11
; CLONE: pMgf10; pMg6Aa; pMgt6-2M
US-08-968-542C-1

Query Match 5.1%; Score 66.4; DB 2; Length 6027;
Best Local Similarity 77.5%; Pred. No. 5.2e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

```
Qy       792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
Db      5606 TCTACGGGCTAGTTTGGGAACCCCATTT-TTCCAAGGGATTTCCATTTTTCCAAGAAAAA 5548
```

```
Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              ||||||||||||   ||   ||||| |||| | ||| ||||||||||||   || |||||
Db       5547 TTAGTTTATTTTTCATTGGAAAAATTGAAATCTCTTGGAAAAATAGAGTTCACATACTAG 5488


RESULT 4
US-09-554-467A-1/c
; Sequence 1, Application US/09554467A
; Patent No. 6639125
; GENERAL INFORMATION:
;  APPLICANT: Myers, Alan M.
;  APPLICANT:  James, Martha G.
;  TITLE OF INVENTION: dull1 Coding for a No. 6639125el Starch Synthase and Uses
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: D6036PCT
;  CURRENT APPLICATION NUMBER: US/09/554,467A
;  CURRENT FILING DATE:  2000-05-12
;  PRIOR APPLICATION NUMBER: PCT/US98/24225
;  PRIOR FILING DATE: 1998-11-12
;  PRIOR APPLICATION NUMBER: US 08/062,102
;  PRIOR FILING DATE: 1997-11-12
;  NUMBER OF SEQ ID NOS: 37
; SEQ ID NO 1
;   LENGTH: 6027
;   TYPE: DNA
;   ORGANISM: maize
;   FEATURE:
;   OTHER INFORMATION: cDNA sequence corresponding to the gene encoding the
;   OTHER INFORMATION: starch synthase enzyme DU1.
US-09-554-467A-1

  Query Match             5.1%;  Score 66.4;  DB 3;  Length 6027;
  Best Local Similarity   77.5%;  Pred. No. 5.2e-06;
  Matches   93;  Conservative    0;  Mismatches   26;  Indels    1;  Gaps    1;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              | || | |||||||||| |||||  ||| || ||| | ||||  |||||| ||| |||||
Db       5606 TCTACGGGCTAGTTTGGGAACCCCATTT-TTCCAAGGGATTTCCATTTTTCCAAGAAAAA 5548

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              ||||||||||||   ||   ||||| |||| | ||| ||||||||||||   || |||||
Db       5547 TTAGTTTATTTTTCATTGGAAAAATTGAAATCTCTTGGAAAAATAGAGTTCACATACTAG 5488


RESULT 5
US-08-410-784A-3/c
; Sequence 3, Application US/08410784A
; Patent No. 5912413
;  GENERAL INFORMATION:
;    APPLICANT:  MYERS, ALAN M.
;    APPLICANT:  JAMES, MARTHA G.
;    TITLE OF INVENTION:  ISOLATION OF SU1, A STARCH DEBRANCHING
;    TITLE OF INVENTION:  ENZYME, THE PRODUCT OF THE MAIZE GENE
;    TITLE OF INVENTION:  SUGARY 1
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weingarten, Schurgin, Gagnebin and Hayes LLP
;      STREET:  Ten Post Office Square
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,784A
;      FILING DATE:  24-MAR-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heine, Ph.D., Holliday C
;      REGISTRATION NUMBER:  34,346
;      REFERENCE/DOCKET NUMBER:  ISU-002XX
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-2290
;      TELEFAX:  617-451-0313
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2523 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
US-08-410-784A-3

  Query Match             4.5%;  Score 58.6;  DB 2;  Length 2523;
  Best Local Similarity   71.4%;  Pred. No. 0.00029;
  Matches  105;  Conservative    0;  Mismatches   39;  Indels    3;  Gaps    2;

Qy        774 AAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTT 833
              | | ||   | |   | |  ||||  |||||||| ||| ||| |||  |  ||||  |||
Db        279 ACATATACATATACATATAATAAGGTCTAGTTTGACAATCCTATTTTATCAAAAG--TTT 222

Qy        834 TGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAA 893
              |    |||||  ||| ||||||||||||||||||||| | ||||||| ||   || ||||||
Db        221 TACATTTTTTCAAATAAATTAGTTTATTTTCTCTTGA-AAAATAGGAATTTCTCAGAAAA 163

Qy        894 ATAGAGTTGCCAGACTAGCCCTAGAAT 920
              ||||||||  || |||||  | || |||
Db        162 ATAGAGTTTACAAACTAGTCTTAAAAT 136


RESULT 6
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 5993827
;  GENERAL INFORMATION:
;    APPLICANT:  Sim, Kim L.
;    APPLICANT:  Chitnis, Chetan
;    APPLICANT:  Miller, Louis H.
;    APPLICANT:  Peterson, David S.
;    APPLICANT:  Su, Xin-zhaun
;    APPLICANT:  Wellems, Thomas E.
;    TITLE OF INVENTION:  BINDING DOMAINS FROM PLASMODIUM VIVAX
;    TITLE OF INVENTION:  AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe Martens Olson & Bear
;      STREET:  620 Newport Center Drive  16th Floor
;      CITY:  Newport Beach
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,826B
;      FILING DATE:  10-SEP-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned
;      REGISTRATION NUMBER:  29,655
```

```
;     REFERENCE/DOCKET NUMBER:  NIH121.001CP1
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (619) 235-8550
;     TELEFAX:  (619) 235-0176
; INFORMATION FOR SEQ ID NO:  13:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  19124 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   HYPOTHETICAL:  NO
;   ANTI-SENSE:  NO
US-08-487-826B-13

  Query Match             4.3%;  Score 57;  DB 2;  Length 19124;
  Best Local Similarity   50.2%;  Pred. No. 0.0014;
  Matches  141;  Conservative    0;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        614 TAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATAT 673
Db      15677 TAATGTTTTTTTTTTCTTCTTTTGTTTTTATTTTTTTATAATCATTTTTTTTTTTATAT 15618

Qy        674 ACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGT 733
Db      15617 AAAATTTTTTTTTAATTTTTTTTTTGATAATCTTTTTCATTTTTTATTCTATCAAAATTTA 15558

Qy        734 CATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTT 793
Db      15557 TATTTTATTATAATTTTTATTATTTTTTAAAAAATTTTTCTCCTTTTTTTTTTTTTTTT 15498

Qy        794 TAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATT 853
Db      15497 TTTTATTTTAATAAATTTTTTTTTTATATTTCATTTTTCTTTTTTTTCATTTTAATAAATTT 15438

Qy        854 AGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAA 894
Db      15437 GTTTTTATATTTCTTTTTTAATATAAATACATATATATAAA 15397

RESULT 7
US-09-662-254B-26/c
; Sequence 26, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:
;  APPLICANT: Moyer, Richard W.
;  APPLICANT:  Li, Yi
;  APPLICANT:  Bawden, Alison Louise
;  TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
;  TITLE OF INVENTION:  Vertebrate Cells
;  FILE REFERENCE: UF-221C1XC1
;  CURRENT APPLICATION NUMBER: US/09/662,254B
;  CURRENT FILING DATE:  2000-09-14
;  PRIOR APPLICATION NUMBER: 09/086,651
;  PRIOR FILING DATE: 1998-05-29
;  PRIOR APPLICATION NUMBER: 60/224,479
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
;   LENGTH: 50000
;   TYPE: DNA
;   ORGANISM: Amsacta moorei entomopoxvirus
US-09-662-254B-26

  Query Match             4.3%;  Score 56.4;  DB 3;  Length 50000;
  Best Local Similarity   45.1%;  Pred. No. 0.0026;
  Matches  257;  Conservative    0;  Mismatches  306;  Indels    7;  Gaps    1;

Qy        315 ATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTAGGGCCTCAGCATAG 374
Db      27972 ATATGTATTAACCAATAAAAAATTAAAAAATGAAATCAAATTTATTACACCTAATAATAT 27913

Qy        375 ATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTGTCCACTATTGGCAC 434
Db      27912 ATTAAGAAAAAGATGGATAATATTATCTGTTGAAAATTTTAAATTATTAATTATGAGATT 27853

Qy        435 TACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGAAAGCTATCATG 494
Db      27852 AAATACTAAATCAGCACATTATATTAGAGAATATTATTTATTTATAGAAGGATTATTGTA 27793

Qy        495 TAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTTCGA 554
Db      27792 TAAATATACAAAATATATTAATAATTACAATAATAATGTTAAAAATAAAGAATTATTAAA 27733

Qy        555 AAGAGACAAATTATGTTACAGTTTACAAACATCTAAGA-------GCGACAAATTATATC 607
Db      27732 ATTAAAAGAAGAAAATAATAATTTACAAAGATTTAATATGAATATGAAAAATTTTATAAA 27673

Qy        608 GAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTT 667
Db      27672 TAATGTTAAAGAAAAAAATAAAAATGGTTATATATATATTGCTACTTCTAAAAGATATGC 27613

Qy        668 TTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTA 727
Db      27612 TTCGATAAATAATTTTAAAGTAGGCAAGACTGATAATTTATCTTCTAGACAATCAAATTT 27553

Qy        728 TAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGA 787
Db      27552 TAATTCATCACACATAGATCAAGATGAATTTTATATATGTTTCTATCAAAAAGTATATAA 27493

Qy        788 TTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAA 847
Db      27492 TATGAGTAAAACAGAAAATTTAATTCACGATTTATTAGAAGATTTTAGAGATAAAAAAAG 27433

Qy        848 AAAATTAGTTTATTTTCTCTTTATAAAATA 877
Db      27432 AAAAGAAATATTTATAATACATTATACATA 27403

RESULT 8
US-09-313-294A-5397
; Sequence 5397, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 5397
;   LENGTH: 279
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700350078H1
;   NAME/KEY: unsure
;   LOCATION: 10, 12, 185-186, 204, 253, 274, 278
;   OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5397

  Query Match             4.3%;  Score 55.8;  DB 3;  Length 279;
  Best Local Similarity   69.0%;  Pred. No. 0.00065;
  Matches   89;  Conservative    0;  Mismatches   39;  Indels    1;  Gaps    1;

Qy        789 TTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAA 848
Db          3 TTCCTAANGNCCTAGTTTGGAAACCCCATTTTCCCCACGGGTTTTTCATTTTCCCAAGGG 62

Qy        849 AAATTAGTTTATTTTCTCTTTAT-AAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGA 907
```

```
Db         63 AAGTTAGAACATTTTCCCCTTGGGAAAATAGAAATCCTTTGGGAAAATCGAGTTCCCAAA 122

Qy        908 CTAGCCCTA 916
              |||||||||
Db        123 CTAGCCCTA 131


RESULT 9
US-09-902-540-1357/c
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
;   LENGTH: 612
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(612)
;   OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

  Query Match             4.2%;  Score 55.6;  DB 3;  Length 612;
  Best Local Similarity   51.4%;  Pred. No. 0.00094;
  Matches  127;  Conservative    0;  Mismatches  120;  Indels    0;  Gaps    0;

Qy        626 TTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCT 685
              ||    | |||| |||||  || || || |||| || || ||||| | |    |||| || |
Db        560 TTTTTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 501

Qy        686 CTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATA 745
              ||     |   | |||| ||    |||| |     | | |||     | |||||| | ||
Db        500 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTA 441

Qy        746 AGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTT 805
              |     | | |   || |  ||       | ||| |  | ||||||     || ||
Db        440 ATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 381

Qy        806 TGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCT 865
              |         | ||| |||    |     |||  |||||          | ||  ||| |||| |
Db        380 TTTTTTTTTTATTTTTTTTTTTATTTATTTTTTTTTTTATTTTTTTANTTTTTTTTTTTTT 321

Qy        866 CTTTATA 872
              || ||
Db        320 TATTTTA 314


RESULT 10
US-08-880-499-1/c
; Sequence 1, Application US/08880499
; Patent No. 6037523
;  GENERAL INFORMATION:
;    APPLICANT:  Albertson, Marc C.
;    APPLICANT:  Fox, Tim W.
;    APPLICANT:  Carl, Garnaat W.
;    APPLICANT:  Huffman, Gary A.
;    APPLICANT:  Kendall, Timmy L.
;    TITLE OF INVENTION:  MALE TISSUE-PREFERRED REGULATORY REGION
;    TITLE OF INVENTION:  AND METHOD OF USING SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PIONEER HI-BRED INTERNATIONAL, INC.
;      STREET:  Darwin Building, 7100 N.W. 62nd Ave., P.O.
;      STREET:  Box 1000
;      CITY:  Johnston
;      STATE:  Iowa
;      COUNTRY:  USA
;      ZIP:  50131
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/880,499
;      FILING DATE:  CONCURRENTLY HEREWITH
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sweeney, Patricia A.
;      REGISTRATION NUMBER:  32,733
;      REFERENCE/DOCKET NUMBER:  0578
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 248-4800
;      TELEFAX:  (515) 248-4844
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1394 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-880-499-1

  Query Match             4.2%;  Score 55.6;  DB 3;  Length 1394;
  Best Local Similarity   55.8%;  Pred. No. 0.0012;
  Matches  106;  Conservative    0;  Mismatches   84;  Indels    0;  Gaps    0;

Qy        761 TAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTC 820
              || |   |  ||  ||  | ||     |    ||  || |||||| ||||||| || |||
Db        950 TACACAGTGATTGTAATTTATTGGGAAAACATTCTAGGGCTAGTCTGGCAACTCTATTTT 891

Qy        821 TTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAA 880
              | | |  |  || |  ||| |  | | |||||  |  | |||||  | |  | |||  | ||
Db        890 TCTAAGTGTTTTCTATTTTATAAAGAGAAAATAAACTAATTTTTTTTGAAAAAATCAAAA 831

Qy        881 AACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAA 940
              |  | | | ||| || ||||||| ||||||  ||  ||     |   || |  ||  ||
Db        830 TTCTTTGAAAGAAACAGGGTTGCCAAACTAGCTCTTAAAAAAATCAGAAAGATTTTTGAA 771

Qy        941 TCACTGTGTA 950
              |   | ||
Db        770 CGAGAATCTA 761


RESULT 11
US-08-880-499-2/c
; Sequence 2, Application US/08880499
; Patent No. 6037523
;  GENERAL INFORMATION:
;    APPLICANT:  Albertson, Marc C.
;    APPLICANT:  Fox, Tim W.
;    APPLICANT:  Carl, Garnaat W.
;    APPLICANT:  Huffman, Gary A.
;    APPLICANT:  Kendall, Timmy L.
;    TITLE OF INVENTION:  MALE TISSUE-PREFERRED REGULATORY REGION
;    TITLE OF INVENTION:  AND METHOD OF USING SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PIONEER HI-BRED INTERNATIONAL, INC.
;      STREET:  Darwin Building, 7100 N.W. 62nd Ave., P.O.
;      STREET:  Box 1000
;      CITY:  Johnston
;      STATE:  Iowa
;      COUNTRY:  USA
;      ZIP:  50131
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/880,499
;      FILING DATE:  CONCURRENTLY HEREWITH
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sweeney, Patricia A.
;      REGISTRATION NUMBER:  32,733
;      REFERENCE/DOCKET NUMBER:  0578
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 248-4800
;      TELEFAX:  (515) 248-4844
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1394 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-880-499-2

  Query Match             4.2%;  Score 55.6;  DB 3;  Length 1394;
  Best Local Similarity   55.8%;  Pred. No. 0.0012;
  Matches  106;  Conservative    0;  Mismatches   84;  Indels    0;  Gaps    0;

Qy        761 TAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTC 820
              || |   |  ||  ||  | ||    |    ||  || |||||| ||||||| || |||
Db        950 TACACAGTGATTGTAATTTATTGGGAAAACATTCTAGGGCTAGTCTGGCAACTCTATTTT 891

Qy        821 TTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAA 880
              | | |  |  || |  ||| |  | | ||||| | |  | ||||| | |  | ||| | ||
Db        890 TCTAAGTGTTTTCTATTTTATAAAGAGAAAATAAACTAATTTTTTTTGAAAAAATCAAAA 831

Qy        881 AACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAA 940
              |  | | | ||| || ||||||| |||||| ||  ||    |    || |  ||  ||
Db        830 TTCTTTGAAAGAAACAGGGTTGCCAAACTAGCTCTTAAAAAAATCAGAAAGATTTTTGAA 771

Qy        941 TCACTGTGTA 950
              |   | ||
Db        770 CGAGAATCTA 761


RESULT 12
US-09-806-708B-23
; Sequence 23, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
;  FILE REFERENCE: 4810-58741
;  CURRENT APPLICATION NUMBER: US/09/806,708B
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/147,133
;  PRIOR FILING DATE: 1999-08-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
;   LENGTH: 1055
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (1)..(1055)
;   OTHER INFORMATION: consensus sequence of A.t. and L.a. FAE1 promoters
US-09-806-708B-23

  Query Match             4.2%;  Score 55.4;  DB 3;  Length 1055;
  Best Local Similarity   22.5%;  Pred. No. 0.0013;
  Matches  182;  Conservative  181;  Mismatches  430;  Indels   15;  Gaps    3;

Qy        292 TCACCTATATTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATAC 351
              :||  |:: :|  : ::||    :|||:: || : : |:: ||:  |:  :       |
Db         37 YCANNTGKRCYARRWGKMTTAYYMTATKWGTTGWAWRTWRWAAMKKTRKWMCSTAMNNAW 96

Qy        352 TTTTGTGTAGGGCCTCAG-CATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCC 410
              ||    :::: |  :  :     |  | :: |::       |:   :| ::: ::: |  :
Db         97 TTMCTARKWRTGTRWWTKTNNNATGTRWWTGYWMTNNNGCSTMTWARRYKTRRWWCYTAM 156

Qy        411 ATGTCTAGTGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAA 470
              ::| ::||     : | |::  ::    :|:|:  :|:  :::|:  |:|:   : : :|
Db        157 WYGASWAGNASTRRTTYTWRWKWMCKRKSARARATRGRARYMRAWYTAWARRTGWTKAMA 216

Qy        471 CTAATCCACTCGAAAGCTATCATGTAATGT---TTAAAGAAACATCTATTAAAACCACGA 527
              : |:|:      | | |::   :|::|::    |||        ||| : :   |  | :
Db        217 YAAWTMNNNNNNAKAACKRATTWGWRAKSNCTCTTAGGTTTKRATCCWAYTCGAGWATKK 276

Qy        528 TCCTCTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATC 587
              :  : | ||: :         ||| :||: | | :::: || |:: | :| :  : :
Db        277 WKTWKTSAAMGMTWNNNNNNNTTTKAAMYAAARWMWWSATTTWWAAAWTSRKTWTYYGR 336

Qy        588 TAAGAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTC 647
              : |       :| ::|:  :: :|    |:    | |: |:|::  ||||  |  |
Db        337 KTAMNNNNGTTCWTRMWAWTWKMWMKTKGTTTWNNNGGRTYTGWTTKKMATTTTAKANNC 396

Qy        648 TTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTT 707
              ||  :.::|     ||  : || :|| :::   |  |:: : | : |: | : : : :
Db        397 TTAAWKWKTCTMNNTTAAKATTYWATCYWKSMTNGTSYRYAAARYTWYAWWTRRYAYANN 456

Qy        708 CCGATTTTATAAAATGACTATA----------AAGTCATTTTTATATAAGAGCACGCAT 756
                ||::|  :  :::|  ||           |  |::|| ::   ::|| : |:  :
Db        457 NNTKTTWKACTWTTYKRCCTTANNTAAWYTKSSANCTSRTTRWKTNCWRAGSKTASMGRA 516

Qy        757 GTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTG 816
                ::  ::    : :: : ::::     :|  :  | |    |    || | |    :
Db        517 YARAYWTGYKWNTAWAYWCWTWYYYRAGAAWTAMYMMTSATCYCATAATTAGTCAGAGGS 576

Qy        817 TTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAAT 876
              | :                | | :: |   || |:|  |  : |       |:|:   |  :
Db        577 TAKGNNNNNNNNCCAATCARWKCTAASAACAMANATTCYCTYAANNATYTWANNATGCW 636

Qy        877 AGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATT 936
                  | ::|       |          |::  |   : | |:|  |::   | | |:
Db        637 NATKTAATMWTNNNNNNAGTWTNNNNNNAKMASAATWYAAAAMTAATKYARTANTTAMAG 696

Qy        937 ACAATCACTGTGTATAATTATTTGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAAT 996
              |:|:  |:|       | | |||     | :   ||||  :              || :
Db        697 AYARAAAYTTRTANNGACTTTTTTNNTTGGMRTNTAAARGWANNNNNNNNNNNNNGACWA 756

Qy        997 CGAGCGAAACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAG 1056
              :  | |          |: | | ||  ||:  : |       | | | ||
Db        757 WRTTTATANCGTNNNNNNNNNNAYATTTNTATTTTWWRTRKANNNNNNNNAAAYYGAAAW 816

Qy       1057 AAATCAGTTTTAAGTCATTGTCCCTGAG 1084
                | : :| :|: : : ||    | ||
Db        817 KNNTTMCWTCKAWMKAWATGAATTTNAG 844


RESULT 13
US-09-004-056-1/c
; Sequence 1, Application US/09004056A
; Patent No. 6566586
; GENERAL INFORMATION:
;  APPLICANT: Calgene LLC
;  TITLE OF INVENTION: Plant Expansin Promoter Sequences
;  FILE REFERENCE: 125
;  CURRENT APPLICATION NUMBER: US/09/004,056A
;  CURRENT FILING DATE: 1998-01-07
;  EARLIER APPLICATION NUMBER: 60034914
;  EARLIER FILING DATE: 1997-07-01
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
```

```
;   LENGTH: 2614
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (930)
;   OTHER INFORMATION: unknown nucleotide
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (947)
;   OTHER INFORMATION: unknown nucleotide
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (956)
;   OTHER INFORMATION: unknown nucleotide
US-09-004-056-1

  Query Match             4.2%;  Score 55.2;  DB 3;  Length 2614;
  Best Local Similarity   54.2%;  Pred. No. 0.0019;
  Matches  155;  Conservative    0;  Mismatches  128;  Indels    3;  Gaps    2;

Qy        611 AGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTA 670
              || | || |   |   |  |  |||||| | || ||| ||  | |||| ||  |   ||||
Db        447 AGTTTAGTTTGGTTATCAATTTTTTTCATATTAATTTTTAGTTTTTTATTTCTAAATTTA 388

Qy        671 TATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAA 730
              | |  |     |  |||    |    |  |||   |  |||| |||||||   || |
Db        387 TGTTGACAAAATGAACTTTATTTTATATATTTTAAATATTATTTGATAAAAT-TTTAAAG 329

Qy        731 AGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTT 790
              |  ||  ||||||      | |   |    ||| |||   ||  |  |||      ||
Db        328 TATTTTTCATATATATATTTCAGGAAAACAATAATTTTCGAATAACGAATTTTTGAGATT 269

Qy        791 TTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAA 850
              ||| ||  | |||| |    |   | ||| |||  | ||||  |  || |||  ||||||
Db        268 TTTAAAATGTTAGTATTTTTAATTTCTTTATTTATATGAATAAT--TTATTTAGAAAAAA 211

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATA 896
              |||||  ||| |     |||| ||   |  |||| | ||| |||||||
Db        210 ATTAGGTTAATGGAAATTTAAAATTTTTAAAATAATTAAAAAAATA 165


RESULT 14
US-09-640-173-53
; Sequence 53, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION:  OVARIAN TUMOR SEQUENCES AND
;  TITLE OF INVENTION:  METHODS OF USE THEREFOR
;  FILE REFERENCE:  210121.484C2
;  CURRENT APPLICATION NUMBER: US/09/640,173
;  CURRENT FILING DATE:  2000-08-15
;  NUMBER OF SEQ ID NOS:  196
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 53
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

  Query Match             4.2%;  Score 55;  DB 3;  Length 396;
  Best Local Similarity   43.8%;  Pred. No. 0.0011;
  Matches  139;  Conservative    0;  Mismatches  178;  Indels    0;  Gaps    0;

Qy        632 TTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACA 691
              ||||| |||||  || || || |||| || || ||||| | |   |||| || | ||
Db         11 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 70

Qy        692 ATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCA 751
              |    |  |||| ||     |||| |     |  | |     |  ||||| | |
Db         71 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 130

Qy        752 CGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAA 811
              | | |   || |  ||       | ||| |  |||||||       |  |||
Db        131 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 190

Qy        812 CCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTAT 871
              | ||| |||       ||||  ||||||  |        ||   || |    | || |
Db        191 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTANNTTNTTTTTTNTTTTNCCTTTNTTTT 250

Qy        872 AAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAAT 931
              ||   | ||||     |  |||| | |     |    |         ||  |     |
Db        251 AATTCANAAAAAGAANAAGAAAANATAANANNNANCNNANNNNNNNNATNNTNCTTNATA 310

Qy        932 AAATTACAATCACTGTG 948
              |      |  | |
Db        311 NTNNTTNNNNNANNGGG 327


RESULT 15
US-09-713-550-53
; Sequence 53, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE:  210121.484C4
;  CURRENT APPLICATION NUMBER: US/09/713,550
;  CURRENT FILING DATE:  2000-11-14
;  NUMBER OF SEQ ID NOS:  205
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 53
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

  Query Match             4.2%;  Score 55;  DB 3;  Length 396;
  Best Local Similarity   43.8%;  Pred. No. 0.0011;
  Matches  139;  Conservative    0;  Mismatches  178;  Indels    0;  Gaps    0;

Qy        632 TTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACA 691
              ||||| |||||  || || || |||| || || ||||| | |   |||| || | ||
Db         11 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 70

Qy        692 ATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCA 751
              |    |  |||| ||     |||| |     |  | |     |  ||||| | |
Db         71 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 130

Qy        752 CGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAA 811
              | | |   || |  ||       | ||| |  |||||||       |  |||
Db        131 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 190

Qy        812 CCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTAT 871
              | ||| |||       ||||  ||||||  |        ||   || |    | || |
Db        191 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTANNTTNTTTTTTNTTTTNCCTTTNTTTT 250

Qy        872 AAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAAT 931
              ||   | ||||     |  |||| | |     |    |         ||  |     |
Db        251 AATTCANAAAAAGAANAAGAAAANATAANANNNANCNNANNNNNNNNATNNTNCTTNATA 310

Qy        932 AAATTACAATCACTGTG 948
              |      |  | |
Db        311 NTNNTTNNNNNANNGGG 327
```

Search completed: March 5, 2006, 22:36:25
Job time : 367.534 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 01:58:03 ; Search time 1368.54 Seconds
(without alignments)
2100.307 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311
Perfect score: 1311
Sequence: 1 ccatggtgtctctatgaaaa.....cgcacatatgtttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	68	5.2	1092	11	US-11-014-071-2	Sequence 2, Appli
	2	57.8	4.4	9347	8	US-10-240-708-36	Sequence 36, Appl
	3	55.6	4.2	254396	7	US-10-330-773-534	Sequence 534, App
	4	54.8	4.2	915	6	US-09-925-065A-54272	Sequence 54272, A
	5	53.6	4.1	488	6	US-09-925-065A-803046	Sequence 803046,
c	6	53.6	4.1	5152	8	US-10-240-708-73	Sequence 73, Appl
	7	53.4	4.1	5562	8	US-10-240-708-63	Sequence 63, Appl
c	8	53.2	4.1	119036	8	US-10-995-561-13314	Sequence 13314, A
	9	52	4.0	488	6	US-09-925-065A-803045	Sequence 803045,
c	10	51.6	3.9	6317	8	US-10-240-708-11	Sequence 11, Appl
c	11	51.6	3.9	355211	7	US-10-330-773-242	Sequence 242, App
	12	51.4	3.9	173602	12	US-11-121-086-25	Sequence 25, Appl
	13	51.2	3.9	687411	7	US-10-330-773-26	Sequence 26, Appl
	14	51	3.9	677	6	US-09-925-065A-667262	Sequence 667262,
	15	51	3.9	677	6	US-09-925-065A-667263	Sequence 667263,
	16	51	3.9	677	6	US-09-925-065A-667265	Sequence 667265,
c	17	50.6	3.9	533	6	US-09-925-065A-13297	Sequence 13297, A
c	18	50.6	3.9	610	6	US-09-925-065A-297109	Sequence 297109,
c	19	50.6	3.9	4330	12	US-11-091-883-182	Sequence 182, App
c	20	50.2	3.8	583	6	US-09-925-065A-206148	Sequence 206148,
	21	49.8	3.8	677	6	US-09-925-065A-667264	Sequence 667264,
	22	49.8	3.8	677	6	US-09-925-065A-667266	Sequence 667266,
c	23	49.8	3.8	4339	8	US-10-909-125-801	Sequence 801, App
c	24	49	3.7	583	6	US-09-925-065A-206150	Sequence 206150,
	25	49	3.7	10467	8	US-10-240-708-2	Sequence 2, Appli
c	26	48.6	3.7	583	6	US-09-925-065A-206149	Sequence 206149,
	27	48.2	3.7	5152	8	US-10-240-708-74	Sequence 74, Appl
	28	48.2	3.7	8467	12	US-11-011-332A-96	Sequence 96, Appl
	29	48.2	3.7	355211	7	US-10-330-773-242	Sequence 242, App
	30	48	3.7	6499	12	US-11-011-332A-76	Sequence 76, Appl
	31	48	3.7	108057	7	US-10-330-773-835	Sequence 835, App
c	32	47.6	3.6	6113	8	US-10-240-708-13	Sequence 13, Appl
c	33	47.4	3.6	173602	12	US-11-121-086-25	Sequence 25, Appl
	34	47.2	3.6	6070	8	US-10-240-708-10	Sequence 10, Appl
	35	47.2	3.6	10619	8	US-10-240-708-3	Sequence 3, Appli
	36	47	3.6	19233	8	US-10-240-708-46	Sequence 46, Appl
	37	46.8	3.6	8537	8	US-10-240-708-42	Sequence 42, Appl
	38	46.8	3.6	8961	8	US-10-240-708-80	Sequence 80, Appl
	39	46.6	3.6	9347	8	US-10-240-708-35	Sequence 35, Appl
c	40	46.6	3.6	86899	7	US-10-330-773-963	Sequence 963, App
	41	46.4	3.5	6499	12	US-11-011-332A-90	Sequence 90, Appl
c	42	46.2	3.5	3219	6	US-09-925-065A-714059	Sequence 714059,
c	43	46.2	3.5	3219	6	US-09-925-065A-714060	Sequence 714060,
	44	46.2	3.5	6499	12	US-11-011-332A-77	Sequence 77, Appl
	45	46.2	3.5	6499	12	US-11-011-332A-91	Sequence 91, Appl

ALIGNMENTS

RESULT 1
US-11-014-071-2
; Sequence 2, Application US/11014071
; Publication No. US20050246796A1
; GENERAL INFORMATION:
; APPLICANT: Cigan, Andrew M.
; APPLICANT: Fox, Timothy W.
; APPLICANT: Hershey, Howard P.
; APPLICANT: Unger, Erica
; APPLICANT: Wu, Yongzhong
; TITLE OF INVENTION: Dominant Gene Suppression Transgenes and
; TITLE OF INVENTION: Methods of Using Same
; FILE REFERENCE: 1554
; CURRENT APPLICATION NUMBER: US/11/014,071
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: 60/530,478
; PRIOR FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/591,975
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)...(1092)
; OTHER INFORMATION: P95
US-11-014-071-2

Query Match 5.2%; Score 68; DB 11; Length 1092;
Best Local Similarity 57.3%; Pred. No. 0.0063;
Matches 142; Conservative 0; Mismatches 105; Indels 1; Gaps 1;

```
Qy      670 ATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATA 729
            | |||  ||   |||  | |  | | ||  |   |  |       || |   ||  ||||
Db      391 ACATAGTTTAGATTCAGTCAGCAAAAAGCTAACATTATGTTTGCTTCTTTTATTCATATA 450

Qy      730 AAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATT 789
              |  ||   ||  ||     |    | |     | |    ||  |  | |||  || | |
Db      451 GTTTTCTTAGCATGAAATTTAAATTCTATATAGTACTGGTTTTTTACGAGCTTACTTAAT 510
```

```
Qy        790 TTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
              ||| | || |||| ||||| |||| ||| || || |||||| ||||| || ||
Db        511 TTTATTAGGGCTAATTTGGTAACCACATTT-TTCCACGGAATTTCAATTTTCCTAAGGAA 569

Qy        850 AATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACT 909
              |||||||| |||||| ||| |||||||||| | | | ||||| ||| ||| |||
Db        570 AATTAGTTAATTTTCGCTTGGGAAAATAGAAATTTCATGGGAAAATGCGGTTCCCAAACT 629

Qy        910 AGCCCTAG 917
              |||| |||
Db        630 AGCCTTAG 637
```

RESULT 2
US-10-240-708-36
; Sequence 36, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/240,708
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 36
; LENGTH: 9347
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-36

Query Match 4.4%; Score 57.8; DB 8; Length 9347;
Best Local Similarity 48.9%; Pred. No. 0.64;
Matches 155; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

```
Qy        547 TATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATAT 606
              | |||      || |  ||  || | |   ||| | | || || |||  |  ||| | |
Db       8530 TTTTTATTTTGATAGGAAGGATTTAAATATTTTTATATATTTAGGAGAAATGAATATTTT 8589

Qy        607 CGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTT 666
              |  | |  || ||  ||   || ||| ||| |  || | | | ||||||| |||||
Db       8590 AAGTATTTTAAATAGGATTTTGGTATATTTTTTTTATTTTTTTAGGTTTTTTGTTTTTGTT 8649

Qy        667 TTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACT 726
              | |  ||  | |||  |   ||    || |  |||||| ||    ||| ||  | |||
Db       8650 TATTAATTTAATTTTATAGTTTTTTTTATATAGATTTTGTTTTTGTTTAATTTATTGATA 8709

Qy        727 ATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTG 786
              | | |  || || | |    | | | |      |   | | |         ||| |
Db       8710 TAATATTTTTTGTTGTTTTTATGAATGTAATGTTTTTTTCGTATTGTTTTTTGTTTTGTT 8769

Qy        787 ATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAA 846
              ||| ||     |  |  |       | |  ||  |||   ||||  |||||| | |
Db       8770 ATTGTTAGTTAAAGGAAAGTTATTGGTTTTTGGATTGTAAATTTTTTTTTTTTTTTTAGA 8829

Qy        847 AAAAATTAGTTTATTTT 863
              || |||   || | ||
Db       8830 TAAGATTTTATTTTGTT 8846
```

RESULT 3
US-10-330-773-534
; Sequence 534, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 534
; LENGTH: 254396
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-534

Query Match 4.2%; Score 55.6; DB 7; Length 254396;
Best Local Similarity 46.2%; Pred. No. 3.3;
Matches 184; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

```
Qy        452 AAAAAAATAACCAAAGTAACTAATCCACTCGAAAGCTATCATGTAATGTTTAAAGAAACA 511
              ||| ||| ||  || | ||  ||    |  ||  |     |  ||     || || |
Db     201251 AAAGAAAAAAGAAAGGAAAAGAAAAGAGGAGAGGGAAGAAAGAAAAGAAAAAAGAAAGGA 201310

Qy        512 TCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTT 571
                 || |  | |||| ||| | |||| | ||   |           |  | ||  ||
Db     201311 AAAGAAAAGAAAAAGATCTTCTGACAAAATAGGCCATTCATTGGGCCAAACATTTTAGTG 201370

Qy        572 ACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGA 631
               | | |    || || || |  | | |||||       |    || | || |  ||||
Db     201371 GCTGGTCTGGAATATTTATGGACTATAAATTTATCTTTAGTTCAAACCATCAGATTCATG 201430

Qy        632 TTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACA 691
              ||| |  |||| ||  ||   |||| | | | | ||||||  |  |  |   ||  |
Db     201431 TTTATTCTTTTAATATTTACCATTTATTAACTATATTTATACTAACATAATGACTTTTTA 201490

Qy        692 ATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCA 751
              | | | | |||| ||         |   ||  |    |||||||  ||||    |||  |
Db     201491 AAACTGGGGTTTTTTTTTTCAGCAACTTAAATAAAGAAAAGTCAGCTTTACTACAGACGA 201550

Qy        752 CGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAA 811
               | ||   | |  |   || | ||    |||| | |  |      |    |||   |
Db     201551 AGAATCAAATGGCATTAGATTTAGAACAACTTCTAAATCATCGTGAGGTATTTTTCCTAT 201610

Qy        812 CCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
                    |  ||  ||  |||||   ||||  |||||
Db     201611 TTTGTCATAATTGTAATCATTTTAGGCTTTTTTAAAAA 201648
```

RESULT 4
US-09-925-065A-54272
; Sequence 54272, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846

```
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 54272
;   LENGTH: 915
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-54272

  Query Match             4.2%;  Score 54.8;  DB 6;  Length 915;
  Best Local Similarity   51.0%;  Pred. No. 1.2;
  Matches  151;  Conservative    1;  Mismatches  143;  Indels    1;  Gaps    1;

Qy        596 ACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTCTTTTTCATTCTTGTTATT 655
Db        347 AAAGATACTATATATATTTATTATATATATTTATATATTTATTATATATAATATATATT 406

Qy        656 TTGTTATTGTTTTTATATACATTTTCT-TCTCTTACAATAGAGTGATTTTCTTCCGATTT 714
Db        407 ATATATAATATATATTATATATATTTTATATATTATATATAATATATATTATATATATAATA 466

Qy        715 TATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCA 774
Db        467 TATAATATATATTATATATTTTATATATWATATATAATATATATTATATATTTTATATTA 526

Qy        775 AAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTT 834
Db        527 TATATAATATATATTTTTATATAATATATATTATATATTTATATATTATATATAATATA 586

Qy        835 GATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGA 890
Db        587 TATTTTTTATATAATATATATTATATATTTTATATATATTTTATATATTATATATA 642
```

```
RESULT 5
US-09-925-065A-803046
; Sequence 803046, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE: 2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 803046
;   LENGTH: 488
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-803046

  Query Match             4.1%;  Score 53.6;  DB 6;  Length 488;
  Best Local Similarity   49.3%;  Pred. No. 1.7;
  Matches  137;  Conservative    1;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        610 AAGGTAAGCTATGACGTTCAGATTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTT 669
Db         37 ATGACATGGTCAGACTTTCTGATTTGTTCTGAACATCCTTTCTTTTTAAACAACCAGTTA 96

Qy        670 ATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATA 729
Db         97 TTTTATTTTAGGACTAAACTTACTACACAGGATTCTTTCTTATATAATAGTATTTCTCTT 156

Qy        730 AAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATT 789
Db        157 TAAACTCTTTTTTACCTCAAAAACTACCTCTTTACAACTTTCTTTACATTTTTTTATTTC 216

Qy        790 TTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
Db        217 CTGGTTCCTTTTACTTTGTTTTATATATAACTTTAAATAAGCTTTGAATTAGACAAAACA 276

Qy        850 AATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTT 887
Db        277 TGTTCTTTTTTTTTTTTTTTTTTAGAAAAAAARACACTT 314
```

```
RESULT 6
US-10-240-708-73/c
; Sequence 73, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT: PIEPENBROCK, Christian
;  APPLICANT: BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION: by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/240,708
;  CURRENT FILING DATE: 2002-10-03
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 73
;   LENGTH: 5152
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-73

  Query Match             4.1%;  Score 53.6;  DB 8;  Length 5152;
  Best Local Similarity   48.2%;  Pred. No. 3;
  Matches  219;  Conservative    0;  Mismatches  224;  Indels   11;  Gaps    2;

Qy        425 CTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGAA 484
Db       4137 CTACAATCCCAACTACTCAAAAATCTCAAACAAAAAAATAACTTAAACCCAAAAATTCAA 4078

Qy        485 AGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAG 544
Db       4077 AACTATCATAAACCGTAATTACGCCACTACATTTCAACCTAAACAAAAAACAAAACCCTA 4018

Qy        545 CATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTAT 604
Db       4017 TCTCAAAAAAAAAAAAAAAAAAAAAAAAAATTTAAAACAATAAAATAACAACGCCATCA 3958

Qy        605 ATCGAAAGGTAAGCTATGACGTTCAGATTTTCTTTTTCATTCTTGTTATTTTGTTATTG 664
Db       3957 TATAACTATTATTATACTTATATTAAAAAATACTTTAAAATTTTTTTATTTTTTATTTTT 3898

Qy        665 TTTTTATATACATTTT-CTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATG 723
Db       3897 TATTTTTAAAATTTTAACTTATTTTAAATTCAAAAAATATATATACAAATTTATTACACG 3838

Qy        724 ACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTT 783
```

```
Db       3837 AACATAAAATCAATTTTCTTTAACACC----------TAAAATTAAAAAATAAAAAATAT 3788

Qy        784 CTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTC 843

Db       3787 TTTCATTAAAACTAATTTCCTTTAACACAATCTTTAATTTATAAAAATTTTCCCATTTTT 3728

Qy        844 AAAAAAAATTAGTTTATTTTCTCTTTATAAAATA 877

Db       3727 AAAAAAAATTCCTAAAATACCTAAATATTAAAAA 3694


RESULT 7
US-10-240-708-63
; Sequence 63, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION:  by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/240,708
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 63
;   LENGTH: 5562
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-63

  Query Match             4.1%;  Score 53.4;  DB 8;  Length 5562;
  Best Local Similarity   52.4%;  Pred. No. 3.3;
  Matches  140;  Conservative    0;  Mismatches  126;  Indels    1;  Gaps    1;

Qy        614 TAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATAT 673

Db       1325 TTATATTTTAAGGTTATTATTTATTTTTTGTTGGGGTGATTTTTTTTTTGTTTTTTTTA 1384

Qy        674 ACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGT 733

Db       1385 TTATTTTAATATTTTATTATTATTTTTTTTTTTTCGGTTAGGAATGATTTTTTTTTGAT 1444

Qy        734 CATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC-TTTCTGATTTTT 792

Db       1445 TATATTTATATTTTTTTATTTTAGTTTAGTTTTTTTTTTTTTTTATTTATTTTTTATTATT 1504

Qy        793 TTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAAT 852

Db       1505 TTTATTTTTTGTTAGTTAGATTTTTTTTTTTTTTATTATTTTTTTGTTTTTTATAGTGTTT 1564

Qy        853 TAGTTTATTTTCTCTTTATAAAATAGA 879

Db       1565 TTTTTTATTTTTTTTTTGTTTAAGAAA 1591


RESULT 8
US-10-995-561-13314/c
; Sequence 13314, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13314
;   LENGTH: 119036
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(119036)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-2
US-10-995-561-13314

  Query Match             4.1%;  Score 53.2;  DB 8;  Length 119036;
  Best Local Similarity   48.4%;  Pred. No. 7.4;
  Matches  177;  Conservative    0;  Mismatches  188;  Indels    1;  Gaps    1;

Qy        619 TATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATT 678

Db      31623 TAAAATCTTATTTTATATATATTTTATATATTTTATATATAAAATCTTATTTTTATATATA 31564

Qy        679 TTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTT 738

Db      31563 TTTTTATATATAAAAAATCATATTTTATATATATTTTATATATTTTATATATATGAAATC 31504

Qy        739 TTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGA 798

Db      31503 TTATTTTATATATTATTTTATATATATTTTATATATATATAAAATCTTATTTTATATATA 31444

Qy        799 GCTAGTTTGGCAACCCTGT-TTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTT 857

Db      31443 TTTTATATATATATAAAATCTTATTTTATATATATTTTATATATATATAAAATCTTATTT 31384

Qy        858 TATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAG 917

Db      31383 TATATATATTTTATATATATATAAAATCTTATTTTATATATATTTTATATATATATAAAA 31324

Qy        918 AATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTGGCCAGCCCCATAAATTA 977

Db      31323 TCTTATTTTATATATATTTTATATATATATAAAATCTTATTTATACAGCCATTTTTCTAA 31264

Qy        978 TTTAAA 983

Db      31263 ATAAAA 31258


RESULT 9
US-09-925-065A-803045
; Sequence 803045, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
```

```
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 803045
;    LENGTH: 488
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-925-065A-803045

  Query Match             4.0%;  Score 52;  DB 6;  Length 488;
  Best Local Similarity   48.9%;  Pred. No. 3.3;
  Matches  136;  Conservative    1;  Mismatches  141;  Indels    0;  Gaps    0;

Qy        610 AAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTT 669
              | |  | | |  ||| ||| ||||| |  |   ||| |||  | |||     |    ||
Db         37 ATGACATGGTCAGACTTTCTGATTTGTTCTGAACATCCTTTCTTTTTAAACAACCAGTTA 96

Qy        670 ATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATA 729
              | ||  ||                | || |  |  ||  |||  || | |||  ||  || |
Db         97 TTTTATTTTAGGACTAAACTTACTACACAGGATTCTTTCTTATATAATAGTATTTCTCTT 156

Qy        730 AAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATT 789
              :  |  |||| ||     |  |  :  || |     ||  | | | |||     |
Db        157 TMAACTCTTTTTTACCTCAAAAACTACCTCTTTACAACTTTCTTTACATTTTTTTATTTC 216

Qy        790 TTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
              |  |       || ||||       | |  |||| ||  |  ||||| ||   ||||| |
Db        217 CTGGTTCCTTTTACTTTGTTTTATATATAACTTTAAATAAGCTTTGAATTAGACAAAACA 276

Qy        850 AATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTT 887
                ||  ||| |||| | ||| || || | || ||||||
Db        277 TGTTCTTTTTTTTTTTTTTTTTAGAAAAAAAGACACTT 314


RESULT 10
US-10-240-708-11/c
; Sequence 11, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION:  by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/240,708
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
;    LENGTH: 6317
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-11

  Query Match             3.9%;  Score 51.6;  DB 8;  Length 6317;
  Best Local Similarity   45.8%;  Pred. No. 7.1;
  Matches  218;  Conservative    0;  Mismatches  254;  Indels    4;  Gaps    1;

Qy        424 ACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGA 483
              || ||    |  |   | || || |  |||||||   | |||  |    ||  |  | |
Db       2030 ACGATAATAAAAAAAAAAAAAAAAAAAAATAAAAAAAACTCAAAAACACTATATAAAAACCA 1971

Qy        484 AAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAA 543
              ||   |  || |||  |||||  |||     ||| ||| ||  ||      |||||||||
Db       1970 AATACAATATATAAACTTTAATTAAAATCTAATTCAAAACAAAATAACTATAAAAAACAT 1911

Qy        544 GCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTA 603
              |    | | | | |.|||  | | |  |  |  || ||  || |   | |||  |
Db       1910 TAAACAATAAATAAAAATAAAAAAATTAATTTTCAAATAATATTAAAAAATTATAAAATT 1851

Qy        604 TATCGAAAGGTAAGCT----ATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGT 659
              |    |||  | |              || |  |  | ||  ||  |  | | ||||   || |
Db       1850 TTCTTAAATATTATAATATAATTATATCTATATAATTTATCCTTACTCTTAAAATATACA 1791

Qy        660 TATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAA 719
              |||     |  | | | |    | |||||  ||||      |  | |||   | ||  |||
Db       1790 TATCAAAATATTTAAAAATAAAATATCTTATAATATTTATAAATACTTTAAAATTATTAA 1731

Qy        720 AATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAAT 779
              |  | ||  |  | |  |  |  |  |  ||    |  |  |  || ||
Db       1730 TCATAAAAAAATTAAAATAATTAAATACATACGACAAAATATTAACAATTATTAAATCTA 1671

Qy        780 CTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTT 839
              ||  | | | | ||  |      | |      | | | ||| |  |   |  |   ||
Db       1670 ATTCATTAATATATTTATTAACAATACTCTTTCAATTTCTCTATATATTTAAAATATTTC 1611

Qy        840 TTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAAT 895
              |  ||||| ||| |    | ||   ||| |||||||      | | |  |||| |
Db       1610 ATAAAAAAATAATAAAAAAACTTAAACTTAATAAAATTACTTAAATTCTAAAAACT 1555


RESULT 11
US-10-330-773-242/c
; Sequence 242, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 242
;    LENGTH: 355211
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)...(355211)
;    OTHER INFORMATION: n = A,T,C or G
US-10-330-773-242

  Query Match             3.9%;  Score 51.6;  DB 7;  Length 355211;
  Best Local Similarity   46.1%;  Pred. No. 18;
  Matches  211;  Conservative    0;  Mismatches  244;  Indels    3;  Gaps    1;

Qy        443 ACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGAAAGCTATCATGTAATGTTT 502
              |||  || |||| ||   ||  ||    |    | | | | | |  |    |
Db      61794 ACACAATATAAATAATATACATAACATATATGTATATTTTATATATGTATATTTTATATA 61735

Qy        503 AAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTTCGAAAGAGACA 562
              |    | || | ||| |   | |    | ||| | |   |  |||| | | |    | |
Db      61734 TATTTTATATATTTTATATATATGTATATTTTATATATAATATATATATTGTATATAATA 61675

Qy        563 AATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATG 622
              |  |||  || | | ||     || ||  |   | | | |||||  | |  ||     |
Db      61674 ATATATAATATATTATA---TTATATATAATATATATAATAAATATATATTA 61618

Qy        623 ACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCT 682
              |  | || | |  | | | | | || |   | |  |||   |  ||||||  || | |
Db      61617 TATATAATATGTATAATATATAATATTTTATATATAATATGTATAATATATATTTTATAT 61558

Qy        683 TCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTAT 742
```

```
Db      61557 ATAATAATATGTACAATATATATTTTATATATAATAATATGTACAATATATATTTTATAT 61498

Qy        743 ATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTA 802

Db      61497 ATAATAATATGTACAATATATATTTTATGTATAATATGTATAATATATATTTTATGTATA 61438

Qy        803 GTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTT 862

Db      61437 ATATATATTTTATGTATAATATATATTTTACGTATATTTTATATATAATATATAATATTT 61378

Qy        863 TCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Db      61377 TATATATAATATATAACATTTTATATATAATATATAAT 61340


RESULT 12
US-11-121-086-25
; Sequence 25, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
;   LENGTH: 173602
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-25

  Query Match             3.9%;  Score 51.4;  DB 12;  Length 173602;
  Best Local Similarity   45.5%;  Pred. No. 17;
  Matches  219;  Conservative    0;  Mismatches  261;  Indels    1;  Gaps    1;

Qy        474 ATCCACTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCT 533

Db     136539 ATATATTATATATTTATATGTTTATGTATATATAAATATATAAT-ATATAAATATATTTA 136597

Qy        534 TAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAG 593

Db     136598 TATATTATATATTTATATAATATTTTATATAATATGTTATATATTTTTATATAATATATT 136657

Qy        594 CGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTA 653

Db     136658 ATATATATTTATATGCTATATTATATAGTATATTATATATAATATATTTTATATATTATA 136717

Qy        654 TTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATT 713

Db     136718 TATATATTTATATAATATTTATATATTATATTTATATAATATGTTTATATATTATTATAT 136777

Qy        714 TTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTC 773

Db     136778 TTATATAATATGTTTATATATTATTATATTTATATAATATGTTTATATATTATTATATTT 136837

Qy        774 AAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTT 833

Db     136838 ATATAATATGTTTATAATATATATAATATATTTATATAATATGTTTATATATTATATATA 136897

Qy        834 TGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAA 893

Db     136898 TAATATATTTATATATAATATGTTTATATATTATATATATAATATATTTATATATTATTATA 136957

Qy        894 ATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAA 953

Db     136958 TATATTTTATAATATATAATATATATAATATAATATATATATTATATATATTATATATAT 137017

Qy        954 T 954

Db     137018 T 137018


RESULT 13
US-10-330-773-26
; Sequence 26, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
;   LENGTH: 687411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(687411)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-26

  Query Match             3.9%;  Score 51.2;  DB 7;  Length 687411;
  Best Local Similarity   46.2%;  Pred. No. 25;
  Matches  170;  Conservative    0;  Mismatches  198;  Indels    0;  Gaps    0;

Qy        531 TCTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAA 590

Db         98 TATATAATAATTAAATTATAGTAAAAATATTTATTTATAATATAAATATAAAATGTATAT 157

Qy        591 GAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTG 650

Db        158 TATATTATAATTATATATAATTATATTGTAATTAAATTATATATAATTTTATATATTATA 217

Qy        651 TTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCG 710

Db        218 TATTATATATATTAAATTATAATACAGCATATTGTAATACAATATATTTTATATTAATAT 277

Qy        711 ATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCG 770

Db        278 AGTATAGTAAAATATATAATATATAATATAATTTCATTATAATAAAATATAATAAAAATA 337

Qy        771 TTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAA 830

Db        338 TTTAAAAATGTATTTTATATCTTACATATATATTCAAATATATACATATATTTATATAAA 397

Qy        831 TTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGA 890

Db        398 TATATACTTAATATTTATATATTAATATATATTTTTATAAATATATATATATTATATAAA 457

Qy        891 AAAATAGA 898

Db        458 TATAAATA 465


RESULT 14
US-09-925-065A-667262
; Sequence 667262, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
```

```
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 667262
;   LENGTH: 677
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-667262

  Query Match             3.9%;  Score 51;  DB 6;  Length 677;
  Best Local Similarity   45.6%;  Pred. No. 5.4;
  Matches  214;  Conservative    1;  Mismatches  251;  Indels    3;  Gaps    1;

Qy        492 ATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTT 551
              || |:|| | |||| |    |||  ||  | |    | ||| | |  |  ||||
Db         31 ATATWATATATAAAATATATAATATATAATATATGTATATTTTATATATAATATATATAA 90

Qy        552 CGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAA 611
              | | |   |  ||   ||| |  ||  |   || ||  |   | | ||| |||  | |
Db         91 ATATATATTATATATAATATATATATTTTATTTATATATAATATATATATTTTATTTATA 150

Qy        612 GGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTAT 671
              |||  |||    || |  | | | |  |  ||  || |  |||| | |   | | ||
Db        151 TATAAAATATATATTTTATTTATATATAATATATATTTATATTTTTTATTTATATATAAT 210

Qy        672 ATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA---ATGACTAT 728
              |||  ||||   | |  || | || |   ||||   |  || | || |   ||  |||
Db        211 ATATATTTATATATTATATATTATATATATTTATATATTATATATTATATATATTTATAT 270

Qy        729 AAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGAT 788
              |   |   |  ||||||     | | ||    ||   | |   |  | |    | |
Db        271 ATTATATATAATATATATATTTATGTATTATATATTATATATATTTATATATGTAATATA 330

Qy        789 TTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAA 848
              | |||     |      | |      |  | | | | | | |  ||  ||  ||  | ||
Db        331 TATTTATGTATTATATATTTATATATTATATATATTATATATATTATATACTTAATATAA 390

Qy        849 AAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGAC 908
              |  || | ||| || | | |||   ||| | |  |  || |   ||| | ||   | |
Db        391 TATATAATATATATTATATATATTGTATATATATAATATATATTTATATATTTTATATAT 450

Qy        909 TAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTAT 957
              |    || | |  | |   |   | || |   |    |||| |||
Db        451 AATATATATATTTATATATAATATATATAAATATATATATTATATATAT 499


RESULT 15
US-09-925-065A-667263
; Sequence 667263, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 667263
;   LENGTH: 677
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-667263

  Query Match             3.9%;  Score 51;  DB 6;  Length 677;
  Best Local Similarity   45.6%;  Pred. No. 5.4;
  Matches  214;  Conservative    1;  Mismatches  251;  Indels    3;  Gaps    1;

Qy        492 ATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTT 551
              || | || | ||||  |    : |||  ||  | |    | ||| | |  |  ||||
Db         31 ATATTATATATAAAATATATWATATATAATATATGTATATTTTATATATAATATATATAA 90

Qy        552 CGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAA 611
              | | |   |  ||   ||| |  ||  |   || ||  |   | | ||| |||  | |
Db         91 ATATATATTATATATAATATATATATTTTATTTATATATAATATATATTTTATTTATA 150

Qy        612 GGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTAT 671
              |||  |||    || |  | | | |  |  ||  || |  |||| | |   | | ||
Db        151 TATAAAATATATATTTTATTTATATATAATATATATTTATATTTTTTATTTATATATAAT 210

Qy        672 ATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA---ATGACTAT 728
              |||  ||||   | |  || | || |   ||||   |  || | || |   ||  |||
Db        211 ATATATTTATATATTATATATTATATATATTTATATATTATATATTATATATATTTATAT 270

Qy        729 AAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGAT 788
              |   |   |  ||||||     | | ||    ||   | |   |  | |    | |
Db        271 ATTATATATAATATATATATTTATGTATTATATATTATATATATTTATATATGTAATATA 330

Qy        789 TTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAA 848
              | |||     |      | |      |  | | | | | | |  ||  ||  ||  | ||
Db        331 TATTTATGTATTATATATTTATATATTATATATATTATATATATTTATATATGTAATATA 390

Qy        849 AAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGAC 908
              |  || | ||| || | | |||   ||| | |  |  || |   ||| | ||   | |
Db        391 TATATAATATATATTATATATATTGTATATATATAATATATATTTATATATTTTATATAT 450

Qy        909 TAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTAT 957
              |    || | |  | |   |   | || |   |    |||| |||
Db        451 AATATATATATTTATATATAATATATATAAATATATATATTATATATAT 499


Search completed: March  6, 2006, 10:32:49
Job time : 1371.54 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 23:29:31 ; Search time 1885.1 Seconds
(without alignments)
5750.985 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311
Perfect score: 1311
Sequence: 1 ccatggtgtctctatgaaaa.....cgcacatatgtttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1311	100.0	1394	8	US-10-713-381-1	Sequence 1, Appli
	2	1311	100.0	1394	8	US-10-713-381-2	Sequence 2, Appli
	3	160.6	12.3	255	8	US-10-713-381-9	Sequence 9, Appli
	4	146	11.1	158	8	US-10-713-381-3	Sequence 3, Appli
	5	80.8	6.2	1261	8	US-10-425-115-134230	Sequence 134230,
	6	74	5.6	320	7	US-10-425-114-23340	Sequence 23340, A
	7	74	5.6	624	7	US-10-425-114-16264	Sequence 16264, A
	8	73	5.6	1326	8	US-10-425-115-141826	Sequence 141826,
	9	72.8	5.6	2445	7	US-10-425-114-32493	Sequence 32493, A
c	10	72.8	5.6	2729	8	US-10-425-115-83293	Sequence 83293, A
c	11	72.4	5.5	1203	8	US-10-425-115-51470	Sequence 51470, A
c	12	71.8	5.5	928	7	US-10-425-114-17816	Sequence 17816, A
c	13	71.8	5.5	1123	8	US-10-425-115-142853	Sequence 142853,
	14	71.8	5.5	2537	7	US-10-425-114-31957	Sequence 31957, A
	15	71.8	5.5	2537	8	US-10-425-115-52216	Sequence 52216, A
	16	71.8	5.5	3607	7	US-10-425-114-31061	Sequence 31061, A
c	17	71.8	5.5	3691	8	US-10-425-115-52219	Sequence 52219, A
c	18	70.2	5.4	960	5	US-10-198-846-6381	Sequence 6381, Ap
	19	69.8	5.3	13400	9	US-10-343-477A-63	Sequence 63, Appl
c	20	69.6	5.3	724	8	US-10-425-115-87756	Sequence 87756, A
c	21	69.2	5.3	527	7	US-10-425-114-4041	Sequence 4041, Ap
	22	68.8	5.2	610	8	US-10-425-115-47165	Sequence 47165, A
c	23	68.4	5.2	1215	7	US-10-425-114-24656	Sequence 24656, A
c	24	68.4	5.2	2863	8	US-10-425-115-75310	Sequence 75310, A
	25	67.8	5.2	512	8	US-10-425-115-49781	Sequence 49781, A
	26	67.8	5.2	928	7	US-10-425-114-17816	Sequence 17816, A
	27	67.8	5.2	1123	8	US-10-425-115-142853	Sequence 142853,
c	28	67.4	5.1	1260	7	US-10-425-114-30881	Sequence 30881, A
c	29	67.4	5.1	1411	8	US-10-425-115-149304	Sequence 149304,
	30	67.2	5.1	2232	7	US-10-425-114-16607	Sequence 16607, A
	31	67.2	5.1	2249	7	US-10-425-114-20264	Sequence 20264, A
c	32	66.4	5.1	6027	7	US-10-634-262-1	Sequence 1, Appli
c	33	66.4	5.1	6027	7	US-10-109-048-1145	Sequence 1145, Ap
c	34	66.4	5.1	6051	8	US-10-425-115-41589	Sequence 41589, A
	35	66.2	5.0	439	3	US-09-960-352-1009	Sequence 1009, Ap
	36	66	5.0	1376	7	US-10-425-114-2073	Sequence 2073, Ap
c	37	65.6	5.0	721	8	US-10-425-115-177935	Sequence 177935,
	38	65.6	5.0	2274	7	US-10-425-114-777	Sequence 777, App
	39	65.6	5.0	2274	7	US-10-425-114-830	Sequence 830, App
	40	65.6	5.0	2632	8	US-10-425-115-162957	Sequence 162957,
c	41	65.6	5.0	8056	8	US-10-473-126-386	Sequence 386, App
	42	65	5.0	1364	8	US-10-425-115-150820	Sequence 150820,
c	43	64.8	4.9	833	8	US-10-425-115-30824	Sequence 30824, A
	44	64.8	4.9	1838	8	US-10-425-115-21012	Sequence 21012, A
c	45	64.2	4.9	581	8	US-10-425-115-48324	Sequence 48324, A

ALIGNMENTS

RESULT 1
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 1311; DB 8; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.2e-275;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
```

```
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 2

```
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1394
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-2

  Query Match            100.0%;  Score 1311;  DB 8;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-275;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
```

```
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Db        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311

Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 3
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 12.3%; Score 160.6; DB 8; Length 255;
Best Local Similarity 97.6%; Pred. No. 7.6e-25;
Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1145 CGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGT 1204

Db          5 CCGCGGATCCCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGT 64

Qy       1205 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 1264

Db         65 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 124

Qy       1265 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311

Db        125 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171
```

RESULT 4
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match 11.1%; Score 146; DB 8; Length 158;
Best Local Similarity 99.4%; Pred. No. 9.1e-22;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 1213

Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy       1214 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 1273

Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120

Qy       1274 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311

Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158
```

RESULT 5
US-10-425-115-134230
; Sequence 134230, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 134230
; LENGTH: 1261
; TYPE: DNA

```
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_53903C.1
US-10-425-115-134230

  Query Match             6.2%;  Score 80.8;  DB 8;  Length 1261;
  Best Local Similarity   78.2%;  Pred. No. 3.9e-07;
  Matches   97;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              | | |  |||||||||  |||| | ||| ||| | ||| |||   |||||| || | |||
Db        999 TCTTAAGGCTAGTTTGAGAACCTTATTTTTTTAAGAGATTTTCTTTTTTTTTAAGATAAA 1058

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              |||||||||||||||||| | |||||||||| | ||| ||||| | ||||||| | |||||
Db       1059 TTAGTTTATTTTCTCTTGAGAAAATAGAAATCTCTTGGAAAATTGGAGTTGCTAAACTAG 1118

Qy        912 CCCT 915
              ||||
Db       1119 CCCT 1122


RESULT 6
US-10-425-114-23340
; Sequence 23340, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 23340
;   LENGTH: 320
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3595-039-G8_FLI
US-10-425-114-23340

  Query Match             5.6%;  Score 74;  DB 7;  Length 320;
  Best Local Similarity   73.1%;  Pred. No. 6e-06;
  Matches   95;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy        791 TTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAA 850
              |||| || |||||||||| || |   ||| ||||||   ||||| |||||   ||| ||
Db         62 TTTTTAGGGCTAGTTTGGAAATCTCATTTTTTTCAAGAGATTTTTATTTTCCTAAAGGAA 121

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
              |||||||||||||   || | |||||||||  | ||| | ||| ||||||| | | || |
Db        122 ATTAGTTTATTTTTCTTTGAGAAAATAGAATTCCCTTGGGAAATTAGAGTTTCTAAACGA 181

Qy        911 GCCCTAGAAT 920
              |||||  | |
Db        182 GCCCTTAATT 191


RESULT 7
US-10-425-114-16264
; Sequence 16264, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 16264
;   LENGTH: 624
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3062-023-G10_FLI
US-10-425-114-16264

  Query Match             5.6%;  Score 74;  DB 7;  Length 624;
  Best Local Similarity   73.1%;  Pred. No. 8.3e-06;
  Matches   95;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy        791 TTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAA 850
              |||| || |||||||||| || |   ||| ||||||   ||||| |||||   ||| ||
Db        374 TTTTTAGGGCTAGTTTGGAAATCTCATTTTTTTCAAGAGATTTTTATTTTCCTAAAGGAA 433

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
              |||||||||||||   || | |||||||||  | ||| | ||| ||||||| | | || |
Db        434 ATTAGTTTATTTTTCTTTGAGAAAATAGAATTCCCTTGGGAAATTAGAGTTTCTAAACGA 493

Qy        911 GCCCTAGAAT 920
              |||||  | |
Db        494 GCCCTTAATT 503


RESULT 8
US-10-425-115-141826
; Sequence 141826, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 141826
;   LENGTH: 1326
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_60829C.1
US-10-425-115-141826

  Query Match             5.6%;  Score 73;  DB 8;  Length 1326;
  Best Local Similarity   69.1%;  Pred. No. 2e-05;
  Matches  114;  Conservative    0;  Mismatches   50;  Indels    1;  Gaps    1;

Qy        756 TGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCT 815
              ||  || | |  | || |    || |  |   |  | ||||  ||| ||||| |||||
Db       1090 TGAAGTTGGTGGTTGTGCTGGTATTTCACCTTTAACTGTAAGGACTAATTTGGGAACCCA 1149

Qy        816 GTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAA 875
              ||| || ||  | ||||| |||||    ||  |||||||||| |||||| ||| | ||||
Db       1150 ATTT-TTCCACTGGATTTTCATTTTCCTAAGGAAAATTAGTTCATTTTCCCTTGAGAAAA 1208

Qy        876 TAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAAT 920
              ||| || | |||||||||| | |||| ||| |||||||||| |||
Db       1209 TAGGAATCCCTTAGAAAAAAATAGTTTCCAAACTAGCCCTAAAAT 1253
```

```
RESULT 9
US-10-425-114-32493
; Sequence 32493, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32493
;   LENGTH: 2445
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLB73402B09_FLI
US-10-425-114-32493

  Query Match             5.6%;  Score 72.8;  DB 7;  Length 2445;
  Best Local Similarity   73.6%;  Pred. No. 2.9e-05;
  Matches  106;  Conservative    0;  Mismatches   37;  Indels    1;  Gaps    1;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTT-TCAAAGAATTTTGATTTTTTCAAAAAAA 850
              | | || |||||||||  |||||| ||| ||   ||||  ||||| ||||||   || ||||
Db       1623 TCTTAGGGCTAGTTTGAGAACCCTTTTTTTTCCCAAAAGATTTTCATTTTTCAAAGAAAA 1682

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
              ||||||| |||||  ||| | ||||||| || | |||| ||||||||  |||   || ||||
Db       1683 ATTAGTTCATTTTTCCTTGAGAAAATAGGAATCCCTTAAAAAAATAGTGTTTTCAAACTA 1742

Qy        911 GCCCTAGAATGTTTTCCCAATAAA 934
              |||||  |   ||||    || ||
Db       1743 GCCCTTAATGTTTTTTTTCATGAA 1766


RESULT 10
US-10-425-115-83293/c
; Sequence 83293, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 83293
;   LENGTH: 2729
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(2729)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_175978C.1
US-10-425-115-83293

  Query Match             5.6%;  Score 72.8;  DB 8;  Length 2729;
  Best Local Similarity   60.7%;  Pred. No. 3.1e-05;
  Matches  136;  Conservative    0;  Mismatches   87;  Indels    1;  Gaps    1;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              |||||| |||||||| | || |||  || |||| |   ||||  |||||| ||  ||||||
Db       2523 TTTAAGGGCTAGTTTAGGAA-CCTCATTTTTTCTAGAGATTTCTATTTTTCCACGAAAAA 2465

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              || || |||| | ||| | |||    || ||  |  ||||||||||||| ||| ||||||
Db       2464 ATATTTCATTTCCCCTTGAAAAATAGAAATACCTTGGAAAAAATAGAGTTTCCAAACTAG 2405

Qy        912 CCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTGGCCAGCCCCAT 971
              || |  ||  || || |||  | || |    |  |    ||||  |       | | |
Db       2404 CCTTCAAAATTTCTCTCAAGTATTTTCTTCTATAGGAAGTAATGGTACACATGGGCAGCT 2345

Qy        972 AAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGA 1015
                | |||    | | ||| || ||  | |  ||| ||||| | |
Db       2344 TGAGTATCATGAGCTAAAATGTAAAAGGGTAAAATCAAATGTTA 2301


RESULT 11
US-10-425-115-51470/c
; Sequence 51470, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 51470
;   LENGTH: 1203
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1203)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_146936C.1
US-10-425-115-51470

  Query Match             5.5%;  Score 72.4;  DB 8;  Length 1203;
  Best Local Similarity   76.1%;  Pred. No. 2.5e-05;
  Matches  102;  Conservative    0;  Mismatches   31;  Indels    1;  Gaps    1;

Qy        787 ATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAA 846
              | ||| ||||| ||||||  || || | | ||| |   |||  ||||  |||||| | |||
Db       1189 ACTTTATTAAGGGCTAGTGGGGGAA-CATATTTTTCCAAAAAGATTTCTATTTTCTTAAA 1131

Qy        847 AAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAG 906
              ||||||||||||||||||||||    |||||| ||| | | |||||||||||||| |||
Db       1130 GAAAATTAGTTTATTTTCTCTTGGAAAAATAAAAATCCCGTAGAAAAATAGAGTTTCCAA 1071

Qy        907 ACTAGCCCTAGAAT 920
              |||||| |||  ||
Db       1070 ACTAGCTCTAAGAT 1057


RESULT 12
US-10-425-114-17816/c
; Sequence 17816, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

```
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 17816
;   LENGTH: 928
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3075-032-A3_FLI
US-10-425-114-17816

  Query Match             5.5%;  Score 71.8;  DB 7;  Length 928;
  Best Local Similarity   78.0%;  Pred. No. 3e-05;
  Matches   99;  Conservative    0;  Mismatches   27;  Indels    1;  Gaps    1;

Qy        790 TTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
              ||||||||||||| |||||||||  |||  || ||  || | ||||  |||||   || |||
Db        804 TTTTTAAGAGCTGGTTTGGCAA-CCTCATTTTTCTAAGGGATTTCTATTTTACTAAGAAA 746

Qy        850 AATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACT 909
              ||||||||  ||||||||||    ||||||  |||   ||| ||||||||||||| ||| | |
Db        745 AATTAGTTAATTTTCTCTTAGAAAAATAAAAATGTCTTGGAAAAATAGAGTTACCAAATT 686

Qy        910 AGCCCTA 916
              | |||||
Db        685 ACCCCTA 679


RESULT 13
US-10-425-115-142853/c
; Sequence 142853, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 142853
;   LENGTH: 1123
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_61761C.1
US-10-425-115-142853

  Query Match             5.5%;  Score 71.8;  DB 8;  Length 1123;
  Best Local Similarity   78.0%;  Pred. No. 3.3e-05;
  Matches   99;  Conservative    0;  Mismatches   27;  Indels    1;  Gaps    1;

Qy        790 TTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
              ||||||||||||| |||||||||  |||  || ||  || | ||||  |||||   || |||
Db        821 TTTTTAAGAGCTGGTTTGGCAA-CCTCATTTTTCTAAGGGATTTCTATTTTACTAAGAAA 763

Qy        850 AATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACT 909
              ||||||||  ||||||||||    ||||||  |||   ||| ||||||||||||| ||| | |
Db        762 AATTAGTTAATTTTCTCTTAGAAAAATAAAAATGTCTTGGAAAAATAGAGTTACCAAATT 703

Qy        910 AGCCCTA 916
              | |||||
Db        702 ACCCCTA 696


RESULT 14
US-10-425-114-31957
; Sequence 31957, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 31957
;   LENGTH: 2537
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLB73238H12_FLI
US-10-425-114-31957

  Query Match             5.5%;  Score 71.8;  DB 7;  Length 2537;
  Best Local Similarity   73.4%;  Pred. No. 4.9e-05;
  Matches  105;  Conservative    0;  Mismatches   37;  Indels    1;  Gaps    1;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              | | || |||||||||  |||||| ||| |  ||||  ||||| ||||||  || |||||
Db       1799 TCTTAGGGCTAGTTTGAGAACCCT-TTTTTCCCAAAAGATTTTCATTTTTCAAAGAAAAA 1857

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              ||||||  |||||  ||| | ||||||| || | |||| ||||||||  |||  || |||||
Db       1858 TTAGTTCATTTTTCCTTGAGAAAATAGGAATCCCTTAAAAAAATAGTGTTTTCAAACTAG 1917

Qy        912 CCCTAGAATGTTTTCCCAATAAA 934
              ||||  |   ||||    || ||
Db       1918 CCCTTAATGTTTTTTTTCATGAA 1940


RESULT 15
US-10-425-115-52216
; Sequence 52216, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 52216
;   LENGTH: 2537
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_147620C.1
US-10-425-115-52216

  Query Match             5.5%;  Score 71.8;  DB 8;  Length 2537;
  Best Local Similarity   73.4%;  Pred. No. 4.9e-05;
  Matches  105;  Conservative    0;  Mismatches   37;  Indels    1;  Gaps    1;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              | | || |||||||||  |||||| ||| |  ||||  ||||| ||||||  || |||||
Db       1799 TCTTAGGGCTAGTTTGAGAACCCT-TTTTTCCCAAAAGATTTTCATTTTTCAAAGAAAAA 1857

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              ||||||  |||||  ||| | ||||||| || | |||| ||||||||  |||  || |||||
Db       1858 TTAGTTCATTTTTCCTTGAGAAAATAGGAATCCCTTAAAAAAATAGTGTTTTCAAACTAG 1917
```

```
Qy       912 CCCTAGAATGTTTTCCCAATAAA 934
             ||||  |   ||||    || ||
Db      1918 CCCTTAATGTTTTTTTTCATGAA 1940
```

Search completed: March 6, 2006, 03:55:14
Job time : 1886.1 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:38:03 ; Search time 685.165 Seconds
(without alignments)
13025.213 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	157	100.0	255	6	AX224402	AX224402 Sequence
	2	157	100.0	1394	6	BD062176	BD062176 Male tiss
	3	157	100.0	1394	6	BD062177	BD062177 Male tiss
	4	157	100.0	1394	6	AX224394	AX224394 Sequence
	5	157	100.0	1394	6	AX224395	AX224395 Sequence
	6	157	100.0	3343	15	AF360356	AF360356 Zea mays
	7	146	93.0	158	6	AX224396	AX224396 Sequence
c	8	54	34.4	110000	15	AP008209_086	Continuation (87 o
c	9	54	34.4	137327	15	AC135206	AC135206 Oryza sat
	10	50	31.8	50	6	AX224398	AX224398 Sequence
	11	40	25.5	40	6	AX224399	AX224399 Sequence
c	12	34.8	22.2	171896	5	BX950854	BX950854 Zebrafish
	13	34.6	22.0	110127	8	AC104134	AC104134 Homo sapi
	14	34.4	21.9	214946	8	AC136896	AC136896 Homo sapi
	15	33.8	21.5	141554	14	CR847532	CR847532 Danio rer
	16	33.8	21.5	214178	5	BX640577	BX640577 Zebrafish
	17	33.4	21.3	211190	14	AC152229	AC152229 Bos tauru
c	18	33.4	21.3	237262	14	AC162712	AC162712 Bos tauru
c	19	33.2	21.1	117764	8	AC011290	AC011290 Homo sapi
c	20	33.2	21.1	162481	14	AC091864	AC091864 Homo sapi
	21	33.2	21.1	177594	8	AC146134	AC146134 Pan trogl
	22	32.8	20.9	82642	8	AC103745	AC103745 Homo sapi
c	23	32.8	20.9	93610	8	AC006602	AC006602 Homo sapi
	24	32.8	20.9	130031	14	AC135264	AC135264 Homo sapi
c	25	32.8	20.9	148598	8	HSBA51C14	AL121875 Human DNA
c	26	32.8	20.9	158316	8	AC135350	AC135350 Homo sapi
	27	32.8	20.9	169771	8	AC087485	AC087485 Homo sapi
	28	32.8	20.9	169867	14	AC019229	AC019229 Homo sapi
c	29	32.8	20.9	193569	8	AC149044	AC149044 Pan trogl
	30	32.6	20.8	53688	14	AC099973	AC099973 Mus muscu
	31	32.6	20.8	110000	14	AC147308_0	AC147308 Pan trogl
	32	32.6	20.8	146080	5	AL953855	AL953855 Zebrafish
	33	32.6	20.8	165845	14	AC127012	AC127012 Rattus no
c	34	32.6	20.8	204190	14	AC140696	AC140696 Rattus no
c	35	32.6	20.8	252534	14	AC097405	AC097405 Rattus no
	36	32.4	20.6	2070	9	BC020027	BC020027 Mus muscu
	37	32.4	20.6	2096	9	BC038066	BC038066 Mus muscu
	38	32.4	20.6	2433	9	BC055469	BC055469 Mus muscu
	39	32.4	20.6	3092	9	BC094506	BC094506 Mus muscu
c	40	32.4	20.6	110000	1	AE017197_05	Continuation (6 of
	41	32.4	20.6	134895	9	AC159811	AC159811 Mus muscu
	42	32.4	20.6	168889	9	AL731779	AL731779 Mouse DNA
c	43	32.4	20.6	182847	14	AC148229	AC148229 Gallus ga
	44	32.4	20.6	221338	5	AC147437	AC147437 Gallus ga
c	45	32.4	20.6	253762	9	AC157998	AC157998 Mus muscu

ALIGNMENTS

RESULT 1
AX224402
LOCUS AX224402 255 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 9 from Patent WO0160997.
ACCESSION AX224402
VERSION AX224402.1 GI:15554644
KEYWORDS .
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 9 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES Location/Qualifiers
source 1. .255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN

Query Match 100.0%; Score 157; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 5.6e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 74

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 134

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db        135 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171
```

```
RESULT 2
BD062176
LOCUS       BD062176               1394 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Male tissue-preferred regulatory region and method of using same.
ACCESSION   BD062176
VERSION     BD062176.1  GI:22607781
KEYWORDS    JP 2001520523-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1394)
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
            Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: JP 2001520523-A 1 30-OCT-2001;
            PIONEER HI BRED INTERNATIONAL INC
COMMENT     PN   JP 2001520523-A/1
            PD   30-OCT-2001
            PF   19-JUN-1998 JP 1999504910
            PR   23-JUN-1997 US    08/880499
            PI   MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
            PI   TIMMY L KENDALL
            PC   C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
            ,C07K14/34,C12Q1/68,
            PC   A01H5/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES            Location/Qualifiers
     source         1. .1394
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
ORIGIN

  Query Match           100.0%;  Score 157;  DB 6;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 5.6e-40;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 3
BD062177
LOCUS       BD062177               1394 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Male tissue-preferred regulatory region and method of using same.
ACCESSION   BD062177
VERSION     BD062177.1  GI:22607782
KEYWORDS    JP 2001520523-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1394)
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
            Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: JP 2001520523-A 2 30-OCT-2001;
            PIONEER HI BRED INTERNATIONAL INC
COMMENT     PN   JP 2001520523-A/2
            PD   30-OCT-2001
            PF   19-JUN-1998 JP 1999504910
            PR   23-JUN-1997 US    08/880499
            PI   MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
            PI   TIMMY L KENDALL
            PC   C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
            ,C07K14/34,C12Q1/68,
            PC   A01H5/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES            Location/Qualifiers
     source         1. .1394
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
ORIGIN

  Query Match           100.0%;  Score 157;  DB 6;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 5.6e-40;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 4
AX224394
LOCUS       AX224394               1394 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 1 from Patent WO0160997.
ACCESSION   AX224394
VERSION     AX224394.1  GI:15554636
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 1 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES            Location/Qualifiers
     source         1. .1394
                    /organism="Zea mays"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:4577"
ORIGIN

  Query Match           100.0%;  Score 157;  DB 6;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 5.6e-40;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
```

```
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 5
AX224395
LOCUS       AX224395               1394 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 2 from Patent WO0160997.
ACCESSION   AX224395
VERSION     AX224395.1  GI:15554637
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 2 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1394
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match           100.0%;  Score 157;  DB 6;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 5.6e-40;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 6
AF360356
LOCUS       AF360356               3343 bp    DNA     linear   PLN 12-MAY-2001
DEFINITION  Zea mays male fertility protein (Ms45) gene, complete cds.
ACCESSION   AF360356
VERSION     AF360356.1  GI:14028756
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3343)
  AUTHORS   Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
  TITLE     Cloning of Ms45, a gene required for male ferility from Zea mays
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3343)
  AUTHORS   Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
            Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
            IA 50131-1004, USA
FEATURES             Location/Qualifiers
     source          1. .3343
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /chromosome="9L"
     gene            <1392. ..>2942
                     /gene="Ms45"
     mRNA            join(<1392. .1768,1898. .2182,2280. .2447,2534. .>2942)
                     /gene="Ms45"
                     /product="male fertility protein"
     CDS             join(1392. .1768,1898. .2182,2280. .2447,2534. .2942)
                     /gene="Ms45"
                     /codon_start=1
                     /product="male fertility protein"
                     /protein_id="AAK52489.1"
                     /db_xref="GI:14028757"
                     /translation="MEKRNLQWRRGRDGIVQYPHLFFAALALALLVADPFGLSPLAEV
                     DYRPVKHELAPYGEVMGSWPRDNASRLRRGRLEFVGEVFGPESIEFDLQGRGPYAGLA
                     DGRVVRWMGEEAGWETFAVMNPDWSEEVCANGVNSTTRKQHEKEEFCGRPLGLRFHGE
                     TGELYVADAYYGLMVVGQSGGVASSVAREADGDPIRFANDLDVHRNGSVFFTDTSMRY
                     SRKDHLNILLEGEGTGRLLRYDPETSGVHVVLKGLVFPNGVQISEDHQFLLFSETTNC
                     RIMRYWLEGPRAGEVEVFANLPGFPDNVRSNGRGQFWVAIDCCRTPAQEVFAKRPWLR
                     TLYFKFPLSLKVLTWKAARRMHTVLALLDGEGRVVEVLEDRGHEVMKLVSEVREVGRK
                     LWIGTVAHNHIATIPYPLED"
ORIGIN

  Query Match           100.0%;  Score 157;  DB 15;  Length 3343;
  Best Local Similarity  100.0%;  Pred. No. 5.5e-40;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 7
AX224396
LOCUS       AX224396               158 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 3 from Patent WO0160997.
ACCESSION   AX224396
VERSION     AX224396.1  GI:15554638
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 3 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .158
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match           93.0%;  Score 146;  DB 6;  Length 158;
  Best Local Similarity  99.4%;  Pred. No. 2.1e-36;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 59
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy         60 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120
```

```
Qy        120 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||||
Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158
```

RESULT 8
AP008209_086/c
WPCOMMENT
Sequence split into 362 fragments LOCUS AP008209 Accession AP008209

Fragment Name	Begin	End
AP008209_000	1	110000
AP008209_001	100001	210000
AP008209_002	200001	310000
AP008209_003	300001	410000
AP008209_004	400001	510000
AP008209_005	500001	610000
AP008209_006	600001	710000
AP008209_007	700001	810000
AP008209_008	800001	910000
AP008209_009	900001	1010000
AP008209_010	1000001	1110000
AP008209_011	1100001	1210000
AP008209_012	1200001	1310000
AP008209_013	1300001	1410000
AP008209_014	1400001	1510000
AP008209_015	1500001	1610000
AP008209_016	1600001	1710000
AP008209_017	1700001	1810000
AP008209_018	1800001	1910000
AP008209_019	1900001	2010000
AP008209_020	2000001	2110000
AP008209_021	2100001	2210000
AP008209_022	2200001	2310000
AP008209_023	2300001	2410000
AP008209_024	2400001	2510000
AP008209_025	2500001	2610000
AP008209_026	2600001	2710000
AP008209_027	2700001	2810000
AP008209_028	2800001	2910000
AP008209_029	2900001	3010000
AP008209_030	3000001	3110000
AP008209_031	3100001	3210000
AP008209_032	3200001	3310000
AP008209_033	3300001	3410000
AP008209_034	3400001	3510000
AP008209_035	3500001	3610000
AP008209_036	3600001	3710000
AP008209_037	3700001	3810000
AP008209_038	3800001	3910000
AP008209_039	3900001	4010000
AP008209_040	4000001	4110000
AP008209_041	4100001	4210000
AP008209_042	4200001	4310000
AP008209_043	4300001	4410000
AP008209_044	4400001	4510000
AP008209_045	4500001	4610000
AP008209_046	4600001	4710000
AP008209_047	4700001	4810000
AP008209_048	4800001	4910000
AP008209_049	4900001	5010000
AP008209_050	5000001	5110000
AP008209_051	5100001	5210000
AP008209_052	5200001	5310000
AP008209_053	5300001	5410000
AP008209_054	5400001	5510000
AP008209_055	5500001	5610000
AP008209_056	5600001	5710000
AP008209_057	5700001	5810000
AP008209_058	5800001	5910000
AP008209_059	5900001	6010000
AP008209_060	6000001	6110000
AP008209_061	6100001	6210000
AP008209_062	6200001	6310000
AP008209_063	6300001	6410000
AP008209_064	6400001	6510000
AP008209_065	6500001	6610000
AP008209_066	6600001	6710000
AP008209_067	6700001	6810000
AP008209_068	6800001	6910000
AP008209_069	6900001	7010000
AP008209_070	7000001	7110000
AP008209_071	7100001	7210000
AP008209_072	7200001	7310000
AP008209_073	7300001	7410000
AP008209_074	7400001	7510000
AP008209_075	7500001	7610000
AP008209_076	7600001	7710000
AP008209_077	7700001	7810000
AP008209_078	7800001	7910000
AP008209_079	7900001	8010000
AP008209_080	8000001	8110000
AP008209_081	8100001	8210000
AP008209_082	8200001	8310000
AP008209_083	8300001	8410000
AP008209_084	8400001	8510000
AP008209_085	8500001	8610000
AP008209_086	8600001	8710000
AP008209_087	8700001	8810000
AP008209_088	8800001	8910000
AP008209_089	8900001	9010000
AP008209_090	9000001	9110000
AP008209_091	9100001	9210000
AP008209_092	9200001	9310000
AP008209_093	9300001	9410000
AP008209_094	9400001	9510000
AP008209_095	9500001	9610000
AP008209_096	9600001	9710000
AP008209_097	9700001	9810000
AP008209_098	9800001	9910000
AP008209_099	9900001	10010000
AP008209_100	10000001	10110000
AP008209_101	10100001	10210000
AP008209_102	10200001	10310000
AP008209_103	10300001	10410000
AP008209_104	10400001	10510000
AP008209_105	10500001	10610000
AP008209_106	10600001	10710000
AP008209_107	10700001	10810000
AP008209_108	10800001	10910000
AP008209_109	10900001	11010000
AP008209_110	11000001	11110000
AP008209_111	11100001	11210000
AP008209_112	11200001	11310000
AP008209_113	11300001	11410000
AP008209_114	11400001	11510000
AP008209_115	11500001	11610000
AP008209_116	11600001	11710000
AP008209_117	11700001	11810000
AP008209_118	11800001	11910000
AP008209_119	11900001	12010000
AP008209_120	12000001	12110000
AP008209_121	12100001	12210000
AP008209_122	12200001	12310000
AP008209_123	12300001	12410000
AP008209_124	12400001	12510000
AP008209_125	12500001	12610000
AP008209_126	12600001	12710000
AP008209_127	12700001	12810000
AP008209_128	12800001	12910000
AP008209_129	12900001	13010000
AP008209_130	13000001	13110000
AP008209_131	13100001	13210000
AP008209_132	13200001	13310000
AP008209_133	13300001	13410000
AP008209_134	13400001	13510000

AP008209_135	13500001	13610000
AP008209_136	13600001	13710000
AP008209_137	13700001	13810000
AP008209_138	13800001	13910000
AP008209_139	13900001	14010000
AP008209_140	14000001	14110000
AP008209_141	14100001	14210000
AP008209_142	14200001	14310000
AP008209_143	14300001	14410000
AP008209_144	14400001	14510000
AP008209_145	14500001	14610000
AP008209_146	14600001	14710000
AP008209_147	14700001	14810000
AP008209_148	14800001	14910000
AP008209_149	14900001	15010000
AP008209_150	15000001	15110000
AP008209_151	15100001	15210000
AP008209_152	15200001	15310000
AP008209_153	15300001	15410000
AP008209_154	15400001	15510000
AP008209_155	15500001	15610000
AP008209_156	15600001	15710000
AP008209_157	15700001	15810000
AP008209_158	15800001	15910000
AP008209_159	15900001	16010000
AP008209_160	16000001	16110000
AP008209_161	16100001	16210000
AP008209_162	16200001	16310000
AP008209_163	16300001	16410000
AP008209_164	16400001	16510000
AP008209_165	16500001	16610000
AP008209_166	16600001	16710000
AP008209_167	16700001	16810000
AP008209_168	16800001	16910000
AP008209_169	16900001	17010000
AP008209_170	17000001	17110000
AP008209_171	17100001	17210000
AP008209_172	17200001	17310000
AP008209_173	17300001	17410000
AP008209_174	17400001	17510000
AP008209_175	17500001	17610000
AP008209_176	17600001	17710000
AP008209_177	17700001	17810000
AP008209_178	17800001	17910000
AP008209_179	17900001	18010000
AP008209_180	18000001	18110000
AP008209_181	18100001	18210000
AP008209_182	18200001	18310000
AP008209_183	18300001	18410000
AP008209_184	18400001	18510000
AP008209_185	18500001	18610000
AP008209_186	18600001	18710000
AP008209_187	18700001	18810000
AP008209_188	18800001	18910000
AP008209_189	18900001	19010000
AP008209_190	19000001	19110000
AP008209_191	19100001	19210000
AP008209_192	19200001	19310000
AP008209_193	19300001	19410000
AP008209_194	19400001	19510000
AP008209_195	19500001	19610000
AP008209_196	19600001	19710000

Query Match 34.4%; Score 54; DB 15; Length 110000;
Best Local Similarity 85.7%; Pred. No. 1.8e-06;
Matches 60; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 123
              ||||| |||||| ||||||| || | ||||||||||||||||||   || ||||||||||
Db      36362 CCTTGGCTATTCCGAACCAACAGAACACCTACTCCCAAACAATCACGCTGACTCATGCAA 36303

Qy        124 CTTCCATGCA 133
              | ||||||||
```

Db 36302 CCTCCATGCA 36293

RESULT 9
AC135206/c
LOCUS AC135206 137327 bp DNA linear PLN 16-APR-2003
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 3 clone OJ1041F02, complete sequence.
ACCESSION AC135206
VERSION AC135206.3 GI:27596977
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 137327)
AUTHORS Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and Collura,K.
TITLE Rice Genomic Sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 137327)
AUTHORS Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C., Currie,J., Collura,K. and Thompson,S.
TITLE Direct Submission
JOURNAL Submitted (09-OCT-2002) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE 3 (bases 1 to 137327)
AUTHORS Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C., Currie,J., Collura,K. and Thompson,S.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-2002) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE 4 (bases 1 to 137327)
AUTHORS Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J., Collura,K. and Thompson,S.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-2003) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE 5 (bases 1 to 137327)
AUTHORS Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T., Saski,C., Henry,D., Thompson,S., Simmons,J., Thurmond,S.K. and Sun,S.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2003) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
COMMENT On Jan 11, 2003 this sequence version replaced gi:24635891.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by more than one plasmid subclone ; and the assembly was confirmed by restriction digest.There are transposons located at area 29427-32874 and at area 19945-33212. Region 70511-70536 is covered by a single subclone. Region 18678-18873 is a double stranded single subclone. Areas 69957-70761, 89836-90416, 98951-99009, 105316-105607, and 133052-133152 are covered by Monsanto only. The nucleotide sequence of this BAC clone was generated by combining Syngenta, Monsanto and Arizona Genomics Institute sequencing data.
FEATURES Location/Qualifiers
source 1. .137327
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OJ1041F02"
/note="(japonica cultivar-group)"
gene complement(638. .9495)
/gene="OJ1041F02.1"
/note="unknown protein"
CDS complement(join(638. .838,1141. .1276,3132. .3244,3326. .3421, 3513. .3727,3812. .3934,4024. .4054,4157. .4216,5088. .5243,

5322. .5411,5484. .5699,5808. .5942,6035. .6538,6622. .6837,
7548. .7593,7679. .7815,8196. .8328,9221. .9495))
/gene="OJ1041F02.1"
/codon_start=1
/product="unknown protein"
/protein_id="AAP06844.1"
/db_xref="GI:29893590"
/translation="MAQSLELLLIQFLMPDNDARRQAEEQIRRLARDPQVVPALVHHL
RTAKTPNVRQLAAVLLRKKITSHWPKLPPHAKASLKQALIDSITIDHSHLVRRASANV
VSIIAKYAVPAGEWPELLPFIFQCSQSPQEDHREVALILFSSLTETIGTTFQSHLNDL
QPILLKCLQDEASSRVRIAALKAVGSFIEYVNDGGDVVKIFRDFVPSILNVSRQCLAN
GEEDVASIAFEIFDELIESPAPLLGDSVRSIVQFSLEVCSNQELEINIRQQAIQIISW
LVKFKASFLKKHKLVIPILQVMCPLLTETADEDGDSDLAADRSAAEVIDTMAINLPRH
VFPPVLEFASVSFRHINPKYREAAVTSLGVVSEGCCEHLKDKLEDCLKVVLEALKDQE
QMVRGAASFALGQFAEHLQPEILSHYESVLPCILNALEDPSDEVKEKSYYALAAFCED
MGENILPYLDPLMCRLVMSLQGSPRNLQETCMSAIGSVAAAAEQAFMPYAEKVLEMMK
GFMVLTNDEDLCARARATEVVGIVAMAVGRARMETILPPFIEAAISGFVLDYSELREY
THGFFSNVAEILDDSFAQYLPHVVPLAFSSCNLDDGSAVDIDDADSVDNGFSGVSSDD
DVNDEPRVRNISWLTIIADILTAIRAIPPAHADVLEKQKDILDTVMNIYIKTMREDDD
KEVVAQACTSLADIVRDCGFAIIEPYITRLADATLILLRQESCCQQVESDGEDDGDID
HDEVLMDAVSDLLPAFAKVMGSYFDPIFTKLFDSLMKFAKSPHPPQDKTMVVATLAEV
AQGMGAPISAYVDKIMPLVLKELASSEATNRRNAAFCVGEMCKNGGAAALKYYGDILH
GLHRLFADSEPDDAVRDNAAGAIARMIMVQPQSIPLNQILPLVPDVINAFAQVVVSPN
ESDEVKTVVAKAVSHLISVYGQQMQPILSALPPAHANALASFANRR"

gene 12372. .15503
/gene="OJ1041F02.2"
/note="unknown protein"

CDS join(12372. .12757,12858. .13025,13130. .13295,13716. .13826,
13982. .14130,14224. .14354,14459. .14558,15197. .15503)
/gene="OJ1041F02.2"
/codon_start=1
/product="unknown protein"
/protein_id="AAP06845.1"
/db_xref="GI:29893591"
/translation="MPPQQGECCRHGWLGVGEAAVGGGGEEPFYVPLRKRLSVDGKAS
TAPRICIWECDGEAGDITCDIVAAPLRRSCSAKAMPPPAPLFRMMTPPPPRPQRGDGE
EARRPGEAIRKGHRSYSLMLNLQLGISYSVGKSSALPFQKLAASDFDPREKVWTRFPP
EGSKFTPPHHSVDFRWKDYCPAVFRHLRKLFGVDPAEYMLAICGNDTLRELASPGKSG
SCFFITQDDRFMIKTVKKSEVKVLIRMLRSYYEHVRQYKSTLLTRFYGTHCIKQAGCP
KIHRRFDLKGSSHGRTIDKTERKIDETTTLKDLDLQYAFRLQRFWYEELMKQIQMDCT
FLETQGIMDYSLLLGVHFRNDYSVSKIGISQHIAFPKSTGKRKSFEGGSSFCELCFVE
SGCKDRDLIDSRKPFIQLGINMPAQAERSSKKILDNFLLNERHLFITPPSGGSCDVYL
FFGIIDILQDYDITKKLEHAYKSFQVNPDYISAVDPKLYSRRFQDFIRRVFIKEQ"

repeat_region 16734. .16878
/note="putative MITE, Gaijin/Gaigin-like"

repeat_region 19263. .19454
/note="putative transposon, CACTA, En/Spm-like"

gene 21819. .24660
/gene="OJ1041F02.3"
/note="hypothetical protein"

CDS join(21819. .22007,22096. .22206,23772. .23996,24523. .24660)
/gene="OJ1041F02.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP06846.1"
/db_xref="GI:29893592"
/translation="MATASDGSGSGGEQRRLLSIPKEGERIIAPTRRPDGTLRKAIRI
RAGYVPQEEVAIYQSKGAQMRKSGPDVPPGYDPALDAKPKTKAAKRNERRKEKRQQAS
TTNDKGKGLHIEDDAGETDNPKDAVDSVTKQISGIAISESLVVATSSTDATDNSKSES
SAPDIDKKIRALKKKIRLAEAQVQGDPENLKPEQLEKMKKIEGWKEELKLLENKSSPA
AS"

gene complement(26563. .30575)
/gene="OJ1041F02.4"
/note="unknown protein"

CDS complement(join(26563. .26772,27082. .27288,27430. .27537,
27881. .27940,28051. .28126,28327. .28484,28616. .28678,
28790. .28874,29560. .30575))
/gene="OJ1041F02.4"
/codon_start=1
/product="unknown protein"
/protein_id="AAP06847.1"
/db_xref="GI:29893593"
/translation="MGPPAAPPSPSSSSGGSSRRRRRLERRNAAKHIGYDASNFCAYP
QSPPAASAPASGSPSLACSPACSLDLTSFRIGGSGDGCRDVQLLCSSLGLSGVDDFAV
PVADWEAHKAGRSSFSSSASTPKPREEPPARDSPVRREVAAEEEPPSLPAPAAAPVLP
AKETPRSVAIEAPAPLLRVDPWEPARPDVRKASGEGGIKGVRPPPVVLKPPPSMVRPA
VCVVESTWDILRSFAPEEDSHAHAPASRSGGDSACQDAGEEEDDAAAVLTLEELRLGE
TSEEFTGTSSLSTTNDDETSSTTTESMFYISPNGRFRRKIRSWNRGMLLGSGSFGTVF
EGISDEGVFFAVKEVCLCDQGSNAQQCIFQLEQEIALLSQFEHENIVQYYGTDKEDSK
LYIFLELVTQGSLASLYQKYRLRDTHVSAYTRQILNGLTYLHERNIVHRDIKCANILV
HANGSVKLADFGLAKEITKFNVLKSCKGTVYWMAPEVVNPKTTYGPEADIWSLGCTVL
EMLTRQLPYPGLEWTQALYRIGKGEPPAIPNCLSRDARDFISQCVKPNPQDRPSAAKL
LEHPFVNRSMRSIRSMRTSSRSNSSILVLLIMWSWPGRTIRYREAILSTAQSSTPTIA
FHLLILSHSLLRGDMSWFSQRLIFSPENRLNVHSSA"

repeat_region 27679. .27832
/note="putative MITE, MITE-adh, type D-like"

repeat_region complement(31668. .31815)

gene complement(35049. .49855)
/gene="OJ1041F02.5"
/note="hypothetical protein"

CDS complement(join(35049. .35184,35866. .36033,36922. .37205,
37280. .38523,40038. .40104,40196. .40287,40406. .40617,
41678. .41760,42229. .42259,42907. .43088,43694. .43753,
44229. .44426,44622. .44676,45877. .45992,46149. .46313,
46447. .46620,47271. .47470,47915. .48019,48200. .48263,
49589. .49855))
/gene="OJ1041F02.5"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP06848.1"
/db_xref="GI:29893594"
/translation="MGGSADPEAPTPTPSPSPSPAKATPSPASADGNRLRRCVQSKLS
WGPPKAGGGGGEAGGAGLPPLAAGDGTPEKVKKRGRPRKSEAGKKPSSNRETTGLEQD
SKDEVILVDESPQKKQRKGRGKNQGAALKVPNRKHCKALESTDGHESCQQLRSSQTQA
VLPQKSPTSVDIDLVTGPSEASPVNDNVDALDNEDKPQLIVDLRSEANIAAEENRRLS
SGKKMHPFFASRKIHKGAGQDILNVEDEDMDSLCAFERDPPLCPVHVLYELEVTMPIH
WSNKWLIADKSFLGTSTTEQNSAEHADPGKHLANFHDKQNKSKFSSQDVIDVDDECLL
ASSSCFHASLFESKQHERVQHELPEVTPKGCQTANLWTDKYRPETAAQVCGNSKHVKF
LNEWLKGWDERGHRNKQNIVTNGSMNGRSCQDGSDTDYSEDASDYENVLLITGPVGCG
KSAAVFACAREQGFNVIELNMPKTVVFPFLPHNKLLAQTLDLKPDKLHDSSHPSIKYI
PTRVFVHMNRNSIISISAHSSSLSVIKLHCLVNTSDMRNGAYVRQKFEEATKSHGLEK
WSQEEIIGLPISNSLDPASGTPGTAEYKQVINKTLILFEDVDTVFDEDRGFISTILKM
VETTKWPIILTSNKKDPPLPHLLAQLVLDFTYPSSAELLSHVDMICKSEGVEITVPQQ
KHIIDAFLGRLNKCLSCPSLLDLDAVHSTVPRIMPWDFPCKLSETIYMEIDKTIVTAE
QKKKQMEVSEFEGLELQIMTPLTKGRSAGKTRKPKKSKLKHGRSADCNDASPCKNDLD
DFHDSPDIPLPSNHQRMRNRRGVVLFAESDDDLADAHAAKDATFTVQEGRLLPQSSEL
PCLYGHGISNIVPESVFFQQSSVPHLHREVISNQLCFPSESRAFEPASSFQNQLESNM
PGSISQICDTFMSQGISCVPESSFMVGGTSASISSDDLLSSLVSNGLSALRNESTYTA
SVVALEDTNKVENQMTDKPQKCMEDEVGETCEAYVELADRNDHASCSITGYQLMDECS
RAESVWLLSGKKNNDSCKVEHVQDTWNRLRQCHPVLPCDMNHNRSVSGALKRVSRVSD
LISESDLMLISCHPFSNDISDPSLTPYTESDGFSYSKQLEMGSIYAQHGLCIFLQDSQ
ATDDGFVDFLQELLFSGTTTTSLGKFVSSGISCGDGSGNISHVKYPTSCISKRRERQA
RLREVLLPVVPPKLSQSLRGPAFVDYLSSMSQISQLENMQLSECKASSKQRRCRQPRH
YLSSGALSLSAEDIGLLAQCSTFSDRRESETIIEQAIS"

repeat_region 39012. .39207
/note="putative MITE, MITE-adh, type M-like"

repeat_region 39681. .39879
/note="putative MITE, MITE-adh, type G-like"

repeat_region 51474. .51836
/note="putative MITE, Castaway-like"

gene complement(59793. .62196)
/gene="OJ1041F02.6"

Query Match 34.4%; Score 54; DB 15; Length 137327;
Best Local Similarity 85.7%; Pred. No. 1.8e-06;
Matches 60; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy        64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 123
             ||||| |||||| ||||||| || | ||||||||||||||||||    || ||||||||||
Db    119499 CCTTGGCTATTCCGAACCAACAGAACACCTACTCCCAAACAATCACGCTGACTCATGCAA 119440

Qy       124 CTTCCATGCA 133
             | ||||||||
Db    119439 CCTCCATGCA 119430
```

```
RESULT 10
AX224398
LOCUS       AX224398                  50 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 5 from Patent WO0160997.
ACCESSION   AX224398
VERSION     AX224398.1  GI:15554640
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 5 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match           31.8%;  Score 50;  DB 6;  Length 50;
  Best Local Similarity  100.0%;  Pred. No. 3.7e-05;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         75 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 50


RESULT 11
AX224399
LOCUS       AX224399                  40 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 6 from Patent WO0160997.
ACCESSION   AX224399
VERSION     AX224399.1  GI:15554641
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 6 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match           25.5%;  Score 40;  DB 6;  Length 40;
  Best Local Similarity  100.0%;  Pred. No. 0.066;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         85 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||||||||||||||||||||||||||||||||||||||
Db          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40


RESULT 12
BX950854/c
LOCUS       BX950854              171896 bp    DNA     linear   VRT 03-FEB-2005
DEFINITION  Zebrafish DNA sequence from clone CH211-125M22, complete sequence.
ACCESSION   BX950854
VERSION     BX950854.12  GI:58190604
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 171896)
  AUTHORS   Pelan,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 25, 2005 this sequence version replaced gi:56368832.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU). For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            CH211-125M22 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .171896
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-125M22"
                     /clone_lib="CHORI-211"
ORIGIN

  Query Match           22.2%;  Score 34.8;  DB 5;  Length 171896;
  Best Local Similarity  52.8%;  Pred. No. 3.1;
  Matches   75;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy         15 TGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATT 74
              |||  |  | |||  ||||  | | ||  | |   ||||  | |  ||  ||||||     | |
Db     158604 TGGGTTTGTCCATACTTGACCCAACGCTGGGTGTGTTAAAACAACCCAGCCTTTTAAAGT 158545

Qy         75 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAA 134
              | | ||       ||   |  || ||| | | ||| | |    || ||  ||  |  ||
Db     158544 AAGTAACATTCATCTAAAAATGCCTAAATATTTCATTTAAAATCTGAAATCTCAGTTTAA 158485

Qy        135 ACACGCACATATGTTTCCTGAA 156
              || | ||||| | |    | ||
Db     158484 ACTCCCACATTTCTCATTTAAA 158463
```

```
RESULT 13
AC104134
LOCUS       AC104134              110127 bp    DNA     linear   PRI 30-APR-2005
DEFINITION  Homo sapiens BAC clone RP11-525L16 from 2, complete sequence.
ACCESSION   AC104134
VERSION     AC104134.4  GI:19551199
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 110127)
  AUTHORS   Doebber,A., Haakenson,W. and Tomlinson,C.
  TITLE     The sequence of Homo sapiens BAC clone RP11-525L16
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 110127)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 110127)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 110127)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 110127)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Mar 20, 2002 this sequence version replaced gi:18677630.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0525L16
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-554H10, 2000 bp overlap;
            the clone sequenced to the right is RP11-450E9, 2000 bp overlap.
            Actual start of this clone is at base position 151996 of
            RP11-554H10.
FEATURES             Location/Qualifiers
     source          1. .110127
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-525L16"
                     /clone_lib="RPCI-11"
     misc_feature    4046. .4309
                     /note="CpG_island (%GC=61.0, o/e=0.83, #CpGs=20)"
     misc_feature    13008. .13484
                     /note="CpG_island (%GC=75.1, o/e=0.69, #CpGs=53)"
     gene            34638. .39571
                     /gene="FLJ25369"
     mRNA            join(34638. .34769,35622. .35658,36392. .36484,39039. .39571)
                     /gene="FLJ25369"
     CDS             join(35632. .35658,36392. .36484,39039. .39461)
                     /gene="FLJ25369"
                     /note="Homo sapiens hypothetical protein FLJ25369
                     (FLJ25369), mRNA.; H_NH0525L16.1
                     This gene was based on gi(22749356)"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAY24330.1"
                     /db_xref="GI:62988943"
                     /translation="MAGVKYPGQDPVDLDIYQSSHMVDYQPYRKHKYSRVTPQEQAKL
                     DAQLRDKEFYRPIPNPNPKLTDGYPAFKRPHMTAKDLGLPGFFPSQEHEATREDERKF
                     TSTCHFTYPASHDLHLAQGDPNQVLQSADFPCLVDPKHQPAAEMAKGYLLLPGCPCLH
                     CHIVKVPILNRWGPLMPFYQ"
     misc_feature    44296. .44562
                     /note="CpG_island (%GC=62.5, o/e=0.89, #CpGs=24)"
     misc_feature    59905. .60161
                     /note="CpG_island (%GC=66.5, o/e=0.76, #CpGs=22)"
     gene            complement(66728. .105650)
                     /gene="EIF2AK3"
     mRNA            complement(join(66728. .67923,68904. .68966,72276. .72377,
                     80861. .81028,84653. .85433,86541. .86690,89505. .89627,
                     93417. .93529,95828. .96048,97969. .98091,98748. .98888,
                     100430. .100592,100805. .101039,103259. .103392,
                     105456. .105650))
                     /gene="EIF2AK3"
     CDS             complement(join(67723. .67923,68904. .68966,72276. .72377,
                     80861. .81028,84653. .85433,86541. .86690,89505. .89627,
                     93417. .93529,95828. .96048,97969. .98091,98748. .98888,
                     100430. .100592,100805. .101039,103259. .103392,
                     105456. .>105650))
                     /gene="EIF2AK3"
                     /note="Homo sapiens eukaryotic translation initiation
                     factor 2-alpha kinase 3 (EIF2AK3), mRNA.; H_NH0525L16.2
                     This gene was based on gi(21361154)
                     Continued from H_NH0450E09.1"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAY24331.1"
                     /db_xref="GI:62988944"
                     /translation="VPGNKMIIPSLDGALFQWDQDRESMETVPFTVESLLESSYKFGD
                     DVVLVGGKSLTTYGLSAYSGKVRYICSALGCRQWDSDEMEQEEDILLLQRTQKTVRAV
                     GPRSGNEKWNFSVGHFELRYIPDMETRAGFIESTFKPNENTEESKIISDVEEQEAAIM
                     DIVIKVSVADWKVMAFSKKGGHLEWEYQFCTPIASAWLLKDGKVIPISLFDDTSYTSN
                     DDVLEDEEDIVEAARGATENSVYLGMYRGQLYLQSSVRISEKFPSSPKALESVTNENA
                     IIPLPTIKWKPLIHSPSRTPVLVGSDEFDKCLSNDKFSHEEYSNGALSILQYPYDNGY
                     YLPYYKRERNKRSTQITVRFLDNPHYNKNIRKKDPVLLLHWWKEIVATILFCIIATTF
                     IVRRLFHPHPHRQRKESETQCQTENKYDSVSGEANDSSWNDIKNSGYISRYLTDFEPI
                     QCLGRGGFGVVFEAKNKVDDCNYAIKRIRLPNRELAREKVMREVKALAKLEHPGIVRY
                     FNAWLEAPPEKWQEKMDEIWLKDESTDWPLSSPSPMDAPSVKIRRMDPFATKEHIEII
```

APSPQRSRSFSVGISCDQTSSSESQFSPLEFSGMDHEDISESVDAAYNLQDSCLTDCD
VEDGTMDGNDEGHSFELCPSEASPYVRSRERTSSSIVPEDSGCDNASSKEEPKTNRLH
IGNHCANKLTAFKPTSSKSSSEATLSISPPRPTTLSLDLTKNTTEKLQPSSPKVYLYI
QMQLCRKENLKDWMNGRCTIEERERSVCLHIFLQIAEAVEFLHSKGLMHRDLKPSNIF
FTMDDVVKVGDFGLVTAMDQDEEEQTVLTPMPAYARHTGQVGTKLYMSPEQIHGNSYS
HKVDIFSLGLILFELLYPFSTQMERVRTLTDVRNLKFPPLFTQKYPCEYVMVQDMLSP
SPMERPEAINIIENAVFEDLDFPGKTVLRQRSRSLSSSGTKHSRQSNNSHSPLPSN"

ORIGIN

Query Match 22.0%; Score 34.6; DB 8; Length 110127;
Best Local Similarity 54.3%; Pred. No. 3.6;
Matches 70; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

```
Qy          3 TGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAA 62
              ||  ||  ||   |  |  |    | | ||  |||||  || | || ||||    | ||
Db      87633 TGGGATTACAAGCGTGAGCCACTGTACCTGGCCAACCTATCTTTTTCTTCCTCTCTTCCA 87692

Qy         63 GCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCA 122
                |||  ||||||||| | ||      | |||||||  |||  ||    ||   ||| |
Db      87693 TCCTCCCCTATTCTTACCCTTTGTTCCATCTACTCCAGAACTCTCACCATTCTCCATACC 87752

Qy        123 ACTTCCATG 131
                || |||| |
Db      87753 ACCTCCAAG 87761
```

RESULT 14
AC136896
LOCUS AC136896 214946 bp DNA linear PRI 21-MAR-2003
DEFINITION Homo sapiens chromosome 15, clone RP11-86001, complete sequence.
ACCESSION AC136896
VERSION AC136896.6 GI:29135656
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 214946)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 15, clone RP11-86001
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 214946)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (09-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 214946)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (23-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 214946)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Mar 21, 2003 this sequence version replaced gi:27877278.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-------------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
Center project name: L27824
Center clone name: 860_O_1

FEATURES Location/Qualifiers
source 1. .214946
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-86001"
/clone_lib="RPCI-11 Human Male BAC"
repeat_region complement(1. .362)
/rpt_family="HAL1"
repeat_region complement(823. .1431)
/rpt_family="L1MC5"
unsure 913. .922
/note="<30 qual SNGL region"
repeat_region complement(2941. .3295)
/rpt_family="MER31B"
unsure complement(3307. .3311)
/note="<30 qual SNGL region"
unsure 3438. .3476
/note="single clone coverage"

```
repeat_region   3546. .3586
                /rpt_family="GA-rich"
repeat_region   3601. .4006
                /rpt_family="MLT1A1"
repeat_region   complement(4008. .4166)
                /rpt_family="MIR"
repeat_region   complement(4580. .4711)
                /rpt_family="L3"
repeat_region   4731. .4751
                /rpt_family="AT_rich"
repeat_region   complement(4758. .4993)
                /rpt_family="MIR"
repeat_region   complement(5336. .5464)
                /rpt_family="L3"
repeat_region   5935. .6412
                /rpt_family="MER39"
repeat_region   complement(6466. .7069)
                /rpt_family="MLT1E"
repeat_region   complement(8633. .8713)
                /rpt_family="MIR"
repeat_region   8768. .9244
                /rpt_family="L1MA2"
repeat_region   complement(9299. .9498)
                /rpt_family="MIR"
repeat_region   10252. .10587
                /rpt_family="MLT1A0"
repeat_region   12232. .12619
                /rpt_family="L1MA5"
repeat_region   complement(13183. .13354)
                /rpt_family="MIR"
repeat_region   14074. .14500
                /rpt_family="L1MB3"
repeat_region   15070. .15201
                /rpt_family="FLAM_C"
repeat_region   complement(16057. .18621)
                /rpt_family="L1MB3"
repeat_region   complement(18678. .19377)
                /rpt_family="L1M1"
repeat_region   complement(19375. .20142)
                /rpt_family="L1MB3"
repeat_region   22615. .22919
                /rpt_family="AluSg"
repeat_region   complement(25101. .25206)
                /rpt_family="HAL1"
repeat_region   complement(25691. .25925)
                /rpt_family="HAL1"
repeat_region   25996. .26033
                /rpt_family="(CAGA)n"
repeat_region   complement(26059. .26582)
                /rpt_family="MLT2B3"
repeat_region   26645. .27455
                /rpt_family="MER21B"
repeat_region   complement(27582. .27833)
                /rpt_family="HAL1"
repeat_region   complement(28185. .28759)
                /rpt_family="LTR68"
repeat_region   complement(29229. .29603)
                /rpt_family="THE1A"
repeat_region   complement(29771. .30036)
                /rpt_family="AluJb"
repeat_region   30056. .30083
                /rpt_family="(TTTA)n"
repeat_region   complement(30090. .30369)
                /rpt_family="AluJo"
repeat_region   31042. .32585
                /rpt_family="L1M4"
repeat_region   32586. .32654
                /rpt_family="(TCTA)n"
repeat_region   32655. .33163
                /rpt_family="L1M4"
repeat_region   33164. .33473
                /rpt_family="AluY"
repeat_region   33474. .36683
                /rpt_family="L1M4"
repeat_region   36696. .36751
                /rpt_family="MLT1D"
repeat_region   36941. .37015
                /rpt_family="MIR"
repeat_region   complement(38001. .38449)
                /rpt_family="L1MD1"
repeat_region   38453. .38855
                /rpt_family="MLT1D"

  Query Match             21.9%;  Score 34.4;  DB 8;  Length 214946;
  Best Local Similarity   54.8%;  Pred. No. 4.2;
  Matches   68;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy          5 TCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGC 64
              ||||||| |||| | | ||  ||  |||  | ||  || ||||||  |  ||| |
Db      35745 TCATCTCTATGGGCAGAAAAAATCATTTGACAAAATTACGCATTCTTCCATGATGTAAAC 35804

Qy         65 CTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
               |    || || |  |||| | || ||     |  ||||  || | ||  ||| ||
Db      35805 ATCCAACATACTAAGAAAAGAAGGTAGCTTTCTTAACCCAATAAATGTAACCTATGAAAA 35864

Qy        125 TTCC 128
                |||
Db      35865 ATCC 35868


RESULT 15
CR847532
LOCUS       CR847532              141554 bp    DNA     linear   HTG 15-JAN-2005
DEFINITION  Danio rerio clone DKEYP-86C9, WORKING DRAFT SEQUENCE, 3 unordered
            pieces.
ACCESSION   CR847532
VERSION     CR847532.4  GI:57863691
KEYWORDS    HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 141554)
  AUTHORS   Pelan,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 15, 2005 this sequence version replaced gi:56309949.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zKp86C9
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 141058 bases at least Q40
            Consensus quality: 141224 bases at least Q30
            Consensus quality: 141289 bases at least Q20
            Insert size: 141354; sum-of-contigs
            Insert size: 152619; 10.9% error; agarose-fp
            Quality coverage: 8.68x in Q20 bases; sum-of-contigs Quality
            coverage: 8.12x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
```

```
            *        1    10378: contig of 10378 bp in length
            *    10379    10478: gap of 100 bp
            *    10479    42327: contig of 31849 bp in length
            *    42328    42427: gap of 100 bp
            *    42428   141554: contig of 99127 bp in length.
FEATURES             Location/Qualifiers
     source          1. .141554
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEYP-86C9"
                     /clone_lib="DanioKeypilot"
     misc_feature    1. .10378
                     /note="assembly_fragment:00017"
     misc_feature    10479. .42327
                     /note="assembly_fragment:00290"
     misc_feature    42428. .141554
                     /note="assembly_fragment:00621.0"
ORIGIN

  Query Match             21.5%;  Score 33.8;  DB 14;  Length 141554;
  Best Local Similarity   52.5%;  Pred. No. 6.6;
  Matches   74;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy          12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
               |||  |    ||||| ||||| |  |  | | | |||| | |  || ||    |||  |
Db       90222 ACAGTGACCCCTACATCCTTGTGCCTCAATGCATATTGTACAAGTGTACAGAGGTTGGGT 90281

Qy          72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
               |    | |  | | ||||   | ||||  || |||  |  | | |||  |||     |
Db       90282 AAGTGGGAATATGGGGATCATTGCTCCTGAAAAATGAAAATAAATCACACAAAGAAACTT 90341

Qy         132 CAAACACGCACATATGTTTCC 152
                || |   || |||   |||
Db       90342 TCAAAAAAAACTTATTAGTCC 90362


Search completed: March  5, 2006, 21:55:45
Job time : 687.29 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:08:23 ; Search time 104.48 Seconds
(without alignments)
10014.946 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	157	100.0	255	5	AAH76340	Aah76340 Z. mays M
	2	157	100.0	1394	2	AAX07408	Aax07408 Zea mays
	3	157	100.0	1394	2	AAX07409	Aax07409 Zea mays
	4	157	100.0	1394	5	AAH76332	Aah76332 Z. mays M
	5	157	100.0	1394	5	AAH76333	Aah76333 Z. mays M
	6	146	93.0	158	5	AAH76334	Aah76334 Z. mays M
	7	54	34.4	2000	11	ACL38730	Acl38730 Rice stre
	8	50	31.8	50	5	AAH76336	Aah76336 Z. mays M
	9	40	25.5	40	5	AAH76337	Aah76337 Z. mays M
	10	32.4	20.6	3267	12	ADK52131	Adk52131 Mouse ato
c	11	31.6	20.1	2352	8	ACA48402	Aca48402 Prokaryot
c	12	31.4	20.0	15734	6	ABN80329	Abn80329 Human che
	13	30.6	19.5	64415	3	AAF22279	Aaf22279 BAC conta
	14	30.6	19.5	79467	9	ADA02717	Ada02717 Mouse Nfa
	15	30.6	19.5	79467	10	ADB72455	Adb72455 Mouse Nfa
	16	30.6	19.5	79467	10	ADE95965	Ade95965 Mouse Nfa
	17	30.4	19.4	52001	12	ADP67076	Adp67076 Human mit
c	18	30.2	19.2	669	3	AAA51436	Aaa51436 A. thalia
	19	30.2	19.2	1266	9	ADB08983	Adb08983 Alloiococ
	20	30.2	19.2	110000	9	ADB12064_07	Continuation (8 of
	21	30	19.1	30	5	AAH76335	Aah76335 Z. mays M
c	22	30	19.1	378	10	ADD33403	Add33403 Mouse mit
	23	30	19.1	396	10	ADD33404	Add33404 Mouse mit
c	24	29.8	19.0	373	4	AAH13328	Aah13328 Human cDN
c	25	29.6	18.9	1185	10	ADF00796	Adf00796 Bacterial
c	26	29.4	18.7	110000	3	AAF22303_2	Continuation (3 of
c	27	29.2	18.6	1475	9	AAL62902	Aal62902 Vernonia
c	28	29.2	18.6	110000	2	AAT42063_06	Continuation (7 of
c	29	28.8	18.3	1500	8	ADA68284	Ada68284 Arabidops
	30	28.6	18.2	756	13	ADS57339	Ads57339 Bacterial
c	31	28.6	18.2	778	4	AAS32892	Aas32892 Human gen
c	32	28.6	18.2	778	4	AAS32893	Aas32893 Human gen
	33	28.6	18.2	1043	10	ADD71973	Add71973 Human uri
	34	28.6	18.2	1522	13	ADW26566	Adw26566 Glycine m
	35	28.6	18.2	1860	6	ABN67847	Abn67847 Streptoco
c	36	28.6	18.2	2160	13	ADX11599	Adx11599 Plant ful
c	37	28.6	18.2	3231	13	ADX13304	Adx13304 Plant ful
c	38	28.6	18.2	51615	11	ACN45162	Acn45162 Human gen
	39	28.6	18.2	110000	13	ABD32806_5	Continuation (6 of
c	40	28.6	18.2	164702	8	ACF62730	Acf62730 Cancer ba
c	41	28.6	18.2	164702	8	ADB20845	Adb20845 MRP1 base
c	42	28.6	18.2	164702	10	ADB87934	Adb87934 Human UGT
c	43	28.6	18.2	164702	10	ADB96917	Adb96917 Human MDR
c	44	28.6	18.2	164702	10	ADB92108	Adb92108 Human MDR
	45	28.4	18.1	1478	3	AAF15843	Aaf15843 Human pro

ALIGNMENTS

RESULT 1
AAH76340
ID AAH76340 standard; DNA; 255 BP.
XX
AC AAH76340;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 promoter fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; promoter; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Example 5; Fig 8; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The

```
CC   present sequence represents a Z. mays Ms45 promoter fragment
XX
SQ   Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 157;  DB 5;  Length 255;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-42;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 74

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 134

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db        135 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171


RESULT 2
AAX07408
ID   AAX07408 standard; DNA; 1394 BP.
XX
AC   AAX07408;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW   Ms45; male; tissue-preferred; regulatory region; plant cells;
KW   plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS   Zea mays.
XX
PN   WO9859061-A1.
XX
PD   30-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US012895.
XX
PR   23-JUN-1997;   97US-00880499.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman GA,  Kendall TL;
XX
DR   WPI; 1999-105628/09.
XX
PT   New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT   - useful in mediating plant fertility, especially hybrid seed production.
XX
PS   Claim 2; Page 22-23; 39pp; English.
XX
CC   The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC   region. It may be used in the construction of a vector for a method of
CC   producing exogenous genes in a male tissue- preferred manner, which is
CC   useful in restoring or conferring fertility, such as in hybrid seed
CC   production. In conferring fertility, a monocot/dicot plant is transformed
CC   with the exogenous nucleotide sequence (a male sterility gene, preferably
CC   Ms45), which encodes a product selected from auxins, rolB and diptheria
CC   toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 157;  DB 2;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-42;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 3
AAX07409
ID   AAX07409 standard; DNA; 1394 BP.
XX
AC   AAX07409;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW   Ms45; male; tissue-preferred; regulatory region; plant cells;
KW   plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS   Zea mays.
XX
PN   WO9859061-A1.
XX
PD   30-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US012895.
XX
PR   23-JUN-1997;   97US-00880499.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman GA,  Kendall TL;
XX
DR   WPI; 1999-105628/09.
XX
PT   New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT   - useful in mediating plant fertility, especially hybrid seed production.
XX
PS   Claim 3; Page 23-24; 39pp; English.
XX
CC   The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC   region. It may be used in the construction of a vector for a method of
CC   producing exogenous genes in a male tissue- preferred manner, which is
CC   useful in restoring or conferring fertility, such as in hybrid seed
CC   production. In conferring fertility, a monocot/dicot plant is transformed
CC   with the exogenous nucleotide sequence (a male sterility gene, preferably
CC   Ms45), which encodes a product selected from auxins, rolB and diptheria
CC   toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 157;  DB 2;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-42;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

```
RESULT 4
AAH76332
ID   AAH76332 standard; DNA; 1394 BP.
XX
AC   AAH76332;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 4; Page 46; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a nucleic acid sequence encoding an Ms45 male
CC   -tissue preferred regulatory region from Z. mays
XX
SQ   Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 157;  DB 5;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-42;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 5
AAH76333
ID   AAH76333 standard; DNA; 1394 BP.
XX
AC   AAH76333;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 4; Page 47; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a nucleic acid sequence encoding an Ms45 male
CC   -tissue preferred regulatory region from Z. mays
XX
SQ   Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 157;  DB 5;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-42;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 6
AAH76334
ID   AAH76334 standard; DNA; 158 BP.
XX
AC   AAH76334;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
```

PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

Query Match 93.0%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 6.1e-39;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 59
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy         60 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120

Qy        120 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158
```

RESULT 7
ACL38730
ID ACL38730 standard; cDNA; 2000 BP.
XX
AC ACL38730;
XX
DT 02-JUN-2005 (first entry)
XX
DE Rice stress-regulated promoter SEQ ID NO:17293.
XX
KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW agriculture.
XX
OS Oryza sativa.
XX
PN WO2003008540-A2.
XX
PD 30-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US019668.
XX
PR 22-JUN-2001; 2001US-0300112P.
PR 24-AUG-2001; 2001US-0314662P.
PR 26-SEP-2001; 2001US-0325277P.
PR 21-NOV-2001; 2001US-0332132P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
PI Moughamer T, Provart N, Ricke D, Zhu T;
XX
DR WPI; 2003-248011/24.
XX
PT New stress-responsive nucleic acid, useful for altering the
PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT stress, salt stress or osmotic stress.
XX
PS Claim 48; SEQ ID NO 17293; 89pp; English.
XX
CC The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention
XX
SQ Sequence 2000 BP; 616 A; 370 C; 355 G; 659 T; 0 U; 0 Other;

Query Match 34.4%; Score 54; DB 11; Length 2000;
Best Local Similarity 85.7%; Pred. No. 1.2e-07;
Matches 60; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 123
              ||||| |||||| ||||||| || | ||||||||||||||||||    || ||||||||||
Db       1819 CCTTGGCTATTCCGAACCAACAGAACACCTACTCCCAAACAATCACGCTGACTCATGCAA 1878

Qy        124 CTTCCATGCA 133
              | |||||||||
Db       1879 CCTCCATGCA 1888
```

RESULT 8
AAH76336
ID AAH76336 standard; DNA; 50 BP.
XX
AC AAH76336;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX

PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -72 to -111 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;

Query Match 31.8%; Score 50; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 50

RESULT 9
AAH76337
ID AAH76337 standard; DNA; 40 BP.
XX
AC AAH76337;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 14; Page 32; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment upstream of the TATA box of a
CC Z. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX
SQ Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;

Query Match 25.5%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
||||||||||||||||||||||||||||||||||||||||
Db 1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40

RESULT 10
ADK52131
ID ADK52131 standard; cDNA; 3267 BP.
XX
AC ADK52131;
XX
DT 20-MAY-2004 (first entry)
XX
DE Mouse atopic-dermatitis/psoriasis-associated EST #4.
XX
KW Mouse; ss; EST; atopic dermatitis; psoriasis; dermatological;
KW anti-inflammatory; antipsoriatic; rash; expressed sequence tag.
XX
OS Mus musculus.
XX
PN WO2004016785-A1.
XX
PD 26-FEB-2004.
XX
PF 06-AUG-2003; 2003WO-JP009999.
XX
PR 06-AUG-2002; 2002JP-00229319.
PR 14-MAY-2003; 2003JP-00136544.
XX
PA (GENO-) GENOX RES INC.
PA (UYJU-) UNIV JUNTENDO.
XX
PI Itoh M, Ogawa K, Shinagawa A, Sudo H, Ogawa H, Ra C;
PI Mitsuishi K;
XX
DR WPI; 2004-214514/20.
XX
PT Detecting atopic dermatitis or psoriasis comprises assaying levels of
PT expression of an indicator gene at a rash site and non-rash site of a
PT person with atopic dermatitis or psoriasis.
XX
PS Claim 20; SEQ ID NO 164; 484pp; Japanese.
XX
CC The invention relates to detecting atopic dermatitis or psoriasis
CC comprising assaying the levels of expression of an indicator gene at a
CC rash site and non-rash site of a person with atopic dermatitis or
CC psoriasis, comparing these levels with those of a healthy person, and
CC determining that if the levels of indicators are higher or lower, then
CC this indicates the disease. Also included are a reagent for detecting
CC atopic dermatitis or psoriasis, a kit for screening for treatments, a
CC transgenic non human vertebrate animal models for the diseases, an agent
CC for inducing the diseases in mice and a DNA chip for assaying for the
CC indicator genes. The method is used for treatment, detection and animal
CC models for research of atopic dermatitis and psoriasis. The present
CC sequence is a Mouse atopic-dermatitis/ psoriasis-associated EST
CC (expressed sequence tag).
XX
SQ Sequence 3267 BP; 1029 A; 658 C; 590 G; 990 T; 0 U; 0 Other;

Query Match 20.6%; Score 32.4; DB 12; Length 3267;
Best Local Similarity 54.1%; Pred. No. 2.5;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 27 TGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAG 86
Db 2774 TTCTTTTTCTCACCTCCCTCTGATTTCTTTGTACCTCACTTTCATTTTGGCCTTTAAGTT 2833

```
Qy          87 GATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATAT 146
               |||     |||||||| || |||| || || |        |   ||| |||||| |  | ||
Db        2834 GTTCTAGATTCCCAAAGAAGCCATTTTGCACTACAGTCAAACATACTAACACCCTGAAAT 2893

Qy         147 GT 148
               ||
Db        2894 GT 2895
```

RESULT 11
ACA48402/c
ID ACA48402 standard; DNA; 2352 BP.
XX
AC ACA48402;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #30059.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Streptococcus mutans.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU44532.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 36272; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2352 BP; 762 A; 401 C; 500 G; 689 T; 0 U; 0 Other;

Query Match 20.1%; Score 31.6; DB 8; Length 2352;
Best Local Similarity 58.5%; Pred. No. 4.2;
Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
Qy          62 AGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGC 121
               | |||  |     |  || | | ||||  | ||||||  || ||||  |  |||    |||
Db         234 ATCCTGACTGTGCCCAAAGCCAAAGGAAGCATACTCCTGAAAAATCTGTGCTACATCTGC 175

Qy         122 AACTTCCATGCAAACACGCACATATGTTTCCTGA 155
               || |||     ||  |||  ||  | ||||||||
Db         174 AATTTCAGCAAAAGAACGTTCAATTTTTTCCTGA 141
```

RESULT 12
ABN80329/c
ID ABN80329 standard; DNA; 15734 BP.
XX
AC ABN80329;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human chemically modified disease associated gene SEQ ID NO 346.
XX
KW Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW antidiabetic; cytostatic; anticonvulsant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200927-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007536.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130908/17.
XX
PT Novel nucleic acid useful for diagnosis and therapy of diseases
PT associated with development genes such as diabetes, comprises a sequence
PT of a segment of chemically pretreated DNA of genes associated with
PT development.
XX
PS Claim 1; SEQ ID NO 346; 27pp; English.
XX
CC The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases in length of a segment of chemically pretreated DNA (II)
CC of genes associated with development selected from 87 genes listed in the
CC specification such as ACCPN, ADFN, or AFD1 and comprising one of 350
CC sequences (ABN79984-ABN80333) or their complements. The invention is
CC useful for the diagnosis or therapy of diseases associated with
CC development genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to

CC histone deacetylation, Currarino syndrome, diseases related with the
CC development of the brain and limb girdle muscular dystrophy and dwarfism.
CC Oligomers specific to each of the genes are useful for detecting the
CC methylation state of all CpG dinucleotides within the 350 sequences or
CC (II) and their complementary sequences, as primer oligonucleotides for
CC the amplification of the 350 sequences, (II) and/or their complements and
CC as oligomer probes for detecting the cytosine methylation state and/or
CC single nucleotide polymorphisms (SNPs). Note: The sequence data for this
CC patent did not form part of the printed specification but is based on
CC sequence information supplied to Derwent by the European Patent Office
XX
SQ Sequence 15734 BP; 3630 A; 744 C; 4234 G; 7126 T; 0 U; 0 Other;

Query Match 20.0%; Score 31.4; DB 6; Length 15734;
Best Local Similarity 59.6%; Pred. No. 9;
Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 52 CCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATC 111
Db 8802 CCTTCTTTCTCCCCTTATCGAATCTTAAACAATCAAAACCCTAAAATAAAACAACCAAAC 8743

Qy 112 TTACTCATGCAACTTCCATGCAAACACGC 140
Db 8742 TACCTAACCCATCTTCCAAATAAACACCC 8714

RESULT 13
AAF22279
ID AAF22279 standard; DNA; 64415 BP.
XX
AC AAF22279;
XX
DT 20-MAR-2001 (first entry)
XX
DE BAC containing repeats from centromeres 1-4 #2.
XX
KW Centromere; michrosome; vector; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200055325-A2.
XX
PD 21-SEP-2000.
XX
PF 17-MAR-2000; 2000WO-US007392.
XX
PR 18-MAR-1999; 99US-0125219P.
PR 01-APR-1999; 99US-0127409P.
PR 18-MAY-1999; 99US-0134770P.
PR 13-SEP-1999; 99US-0153584P.
PR 17-SEP-1999; 99US-0154603P.
PR 16-DEC-1999; 99US-0172493P.
XX
PA (UYCH-) UNIV CHICAGO.
XX
PI Preuss D, Copenhaver G, Keith K;
XX
DR WPI; 2000-587529/55.
XX
PT Recombinant DNA construct comprising a plant centromere, useful for
PT producing stably inherited michrosomes which can serve as vectors for the
PT construction of transgenic plant and animal cells.
XX
PS Claim 102; Page 321-335; 1449pp; English.
XX
CC The present invention relates to a recombinant DNA construct of a plant
CC (Arabidopsis thaliana) centromere. The constructs are useful for
CC producing stably inherited michrosomes which can serve as vectors for the
CC construction of transgenic plant and animal cells expressing selected
CC proteins such as hormones, enzymes, interleukins, clotting factors,
CC cytokines, antibodies, and growth factors
XX
SQ Sequence 64415 BP; 18698 A; 13554 C; 13083 G; 18980 T; 0 U; 100 Other;

Query Match 19.5%; Score 30.6; DB 3; Length 64415;
Best Local Similarity 58.1%; Pred. No. 26;
Matches 54; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 65 CTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
Db 7117 CTTGCTTCTTTAAGTATTATAGTATATTTGCTCCTAAACACTAAACCTAAACCCTACACC 7176

Qy 125 TTCCATGCAAACACGCACATATGTTTCCTGAAC 157
Db 7177 TTAAATCCCAAACCCTAAATCTAATTCCTTAAC 7209

RESULT 14
ADA02717
ID ADA02717 standard; DNA; 79467 BP.
XX
AC ADA02717;
XX
DT 06-NOV-2003 (first entry)
XX
DE Mouse Nfatc1 carcinoma associated gene, SEQ ID NO:1235.
XX
KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX
OS Mus sp.
XX
PN WO2003057146-A2.
XX
PD 17-JUL-2003.
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
PR 26-DEC-2001; 2001US-00035832.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-587068/55.
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX
PS Claim 1; SEQ ID NO 1235; 245pp; English.
XX
CC The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed murine CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;

Query Match 19.5%; Score 30.6; DB 9; Length 79467;
Best Local Similarity 56.4%; Pred. No. 28;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
Qy         51 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 110
              ||||    |  | ||||||  |  |||| | ||||  ||  | || || || |||  |
Db      53779 TCCACTAACTGATCCTTGCTTCCTCTGGAACAAGCTGAGTCTTAGTCACACACACACACA 53838

Qy        111 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTC 151
              |  || ||  ||    |||  || ||||   | | ||  ||
Db      53839 CACACACACACACACCACACACACACAAGCTTGTGGCTC 53879
```

RESULT 15
ADB72455
ID ADB72455 standard; DNA; 79467 BP.
XX
AC ADB72455;
XX
DT 04-DEC-2003 (first entry)
XX
DE Mouse Nfatc1 gene.
XX
KW mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Mus sp.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
XX
PR 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2003-239337/23.
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS Claim 1; SEQ ID NO 283; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a mouse gene of the invention.
XX
SQ Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;

Query Match 19.5%; Score 30.6; DB 10; Length 79467;
Best Local Similarity 56.4%; Pred. No. 28;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
Qy         51 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 110
              ||||    |  | ||||||  |  |||| | ||||  ||  | || || || |||  |
Db      53779 TCCACTAACTGATCCTTGCTTCCTCTGGAACAAGCTGAGTCTTAGTCACACACACACACA 53838

Qy        111 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTC 151
              |  || ||  ||    |||  || ||||   | | ||  ||
Db      53839 CACACACACACACACCACACACACACAAGCTTGTGGCTC 53879
```

Search completed: March 5, 2006, 18:11:42
Job time : 105.48 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 18:11:59 ; Search time 741.207 Seconds
(without alignments)
9910.279 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
c 1	157	100.0	687	9	CC656939	CC656939 OGWDQ20TV
2	157	100.0	915	10	CG224225	CG224225 OG1AG08TV
3	157	100.0	963	9	CC656933	CC656933 OGWDQ20TM
4	63	40.1	702	10	CW324514	CW324514 104_819_1
5	47	29.9	296	10	CW445575	CW445575 fsbb001f1
c 6	37.8	24.1	636	10	CE588104	CE588104 tigr-gss-
c 7	37.8	24.1	643	7	CO689495	CO689495 DG11-25n6
c 8	37.8	24.1	1085	10	AG366773	AG366773 Mus muscu
9	35.6	22.7	874	10	CZ727957	CZ727957 OC__Ba005
10	34	21.7	700	2	BB505306	BB505306 BB505306
c 11	33.8	21.5	734	10	CW871670	CW871670 she2h63-4
12	33.6	21.4	1677	10	CL026582	CL026582 CH216-23G
13	32.8	20.9	440	1	AI744861	AI744861 tr16b06.x
c 14	32.6	20.8	648	8	CX446438	CX446438 JGI_XZG22
c 15	32.6	20.8	740	10	CW558661	CW558661 OA_ABa008
c 16	32.6	20.8	758	10	CW597425	CW597425 OA_ABa013
17	32.6	20.8	775	8	CX424281	CX424281 JGI_XZG64
18	32.6	20.8	804	7	CR575123	CR575123 CR575123
c 19	32.6	20.8	843	10	CW599444	CW599444 OA_ABa013
20	32.6	20.8	864	8	CX404228	CX404228 JGI_XZT35
21	32.6	20.8	868	8	DN096073	DN096073 JGI_CABE6
c 22	32.6	20.8	1101	10	CNS003FV	AL064657 Drosophil
23	32.4	20.6	347	2	BE690384	BE690384 uw66f12.y
c 24	32.4	20.6	483	1	AJ746802	AJ746802 AJ746802
c 25	32.4	20.6	497	1	AW990724	AW990724 uf10b11.x
c 26	32.4	20.6	498	2	BE632212	BE632212 uu11f01.x
c 27	32.4	20.6	500	2	BF225639	BF225639 uy44e02.x
28	32.4	20.6	527	2	BE448824	BE448824 ut88c10.y
29	32.4	20.6	528	7	CN243061	CN243061 EST008934
30	32.4	20.6	539	5	BX528623	BX528623 BX528623
31	32.4	20.6	542	1	AA673664	AA673664 vo57h06.r
c 32	32.4	20.6	543	2	BE133867	BE133867 ug13b03.x
c 33	32.4	20.6	545	2	BF020472	BF020472 uw66f12.x
34	32.4	20.6	547	5	BX517541	BX517541 BX517541
c 35	32.4	20.6	570	2	BE134144	BE134144 ug16f12.x
36	32.4	20.6	621	1	BB004373	BB004373 BB004373
37	32.4	20.6	742	3	BI453933	BI453933 603174018
38	32.4	20.6	859	10	CNS004YY	AL055406 Drosophil
c 39	32.4	20.6	914	2	BF682258	BF682258 602117694
c 40	32.4	20.6	1039	11	CNS055LA	AL322183 Tetraodon
41	32.4	20.6	2194	4	BC048853	BC048853 Mus muscu
42	32.4	20.6	2454	4	AK028832	AK028832 Mus muscu
43	32.2	20.5	432	9	BH752563	BH752563 SALK_0188
44	32.2	20.5	960	10	DU054394	DU054394 91884 Tom
c 45	32.2	20.5	979	10	CZ954687	CZ954687 268372 To

ALIGNMENTS

RESULT 1
CC656939/c
LOCUS CC656939 687 bp DNA linear GSS 19-JUN-2003
DEFINITION OGWDQ20TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15, genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 687)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGWDQ20TM
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: methylation filtered.
FEATURES Location/Qualifiers
source 1. .687
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0554D15"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 100.0%; Score 157; DB 9; Length 687;
Best Local Similarity 100.0%; Pred. No. 2.2e-39;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 362

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 302

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
                |||||||||||||||||||||||||||||||||||||
Db        301 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 265


RESULT 2
CG224225
LOCUS       CG224225                915 bp    DNA     linear   GSS 22-AUG-2003
DEFINITION  OG1AG08TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0716B15,
            genomic survey sequence.
ACCESSION   CG224225
VERSION     CG224225.1  GI:34124113
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1AG08TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0716B15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             100.0%;  Score 157;  DB 10;  Length 915;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-39;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 593

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 653

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
                |||||||||||||||||||||||||||||||||||||
Db        654 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 690


RESULT 3
CC656933
LOCUS       CC656933                963 bp    DNA     linear   GSS 19-JUN-2003
DEFINITION  OGWDQ20TM ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15,
            genomic survey sequence.
ACCESSION   CC656933
VERSION     CC656933.1  GI:32060225
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 963)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OGWDQ20TV
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .963
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0554D15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             100.0%;  Score 157;  DB 9;  Length 963;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-39;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        752 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 811

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        812 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 871

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
                |||||||||||||||||||||||||||||||||||||
Db        872 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 908


RESULT 4
CW324514
LOCUS       CW324514                702 bp    DNA     linear   GSS 31-OCT-2004
DEFINITION  104_819_11477203_148_35910_078 Sorghum methylation filtered library
            (LibID: 104) Sorghum bicolor genomic clone 11477203, genomic survey
            sequence.
ACCESSION   CW324514
VERSION     CW324514.1  GI:55040702
KEYWORDS    GSS.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
            Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
            McMenamy,J., Smith,M., Holeman,H., Roe,B.A, Wiley,G., Korf,I.F.,
            Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
            Martienssen,R.A.
```

```
  TITLE     Sorghum genome sequencing by methylation filtration
  JOURNAL   PLoS Biol. 3 (1), e13 (2005)
   PUBMED   15660154
COMMENT     Contact: Bedell JA
            Orion Genomics, LLC
            4041 Forest Park Ave, St. Louis, MO 63108, USA
            Tel: 314 615 6979
            Fax: 314 615 5975
            Email: jbedell@oriongenomics.com
            Plate: 819  row: c  column: 19
            Seq primer: SWfor Forward
            Class: methylation filtered
            High quality sequence stop: 702.
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="Sorghum bicolor"
                     /mol_type="genomic DNA"
                     /cultivar="ATx623"
                     /db_xref="taxon:4558"
                     /clone="11477203"
                     /clone_lib="Sorghum methylation filtered library (LibID:
                     104)"
                     /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
                     prepared from purified nuclei was randomly sheared,
                     end-repaired, size fractionated to enrich for the 0.5 to 5
                     kb fraction, ligated into HincII-digested pBCSK(-) vector
                     and electroporated into E. coli cells. This is a
                     methylation filtered library."
ORIGIN

  Query Match            40.1%;  Score 63;  DB 10;  Length 702;
  Best Local Similarity  93.0%;  Pred. No. 3.9e-09;
  Matches   66;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 123
              || || |||||||||||||||||| |||||||||||||||||||| |||| |||||||
Db         53 CCCTGGCTATTCTGAACCAAGAGGACACCTACTCCCAAACAATCCATGTTACCCATGCAA 112

Qy        124 CTTCCATGCAA 134
              ||||||||||
Db        113 CTTCCATGCAA 123


RESULT 5
CW445575
LOCUS       CW445575                296 bp    DNA     linear   GSS 02-NOV-2004
DEFINITION  fsbb001f170m16k0 Sorghum methylation filtered library (LibID: 104)
            Sorghum bicolor genomic clone fsbb001f170m16, genomic survey
            sequence.
ACCESSION   CW445575
VERSION     CW445575.1  GI:55193536
KEYWORDS    GSS.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 296)
  AUTHORS   Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
            Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
            McMenamy,J., Smith,M., Holeman,H., Roe,B.A, Wiley,G., Korf,I.F.,
            Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
            Martienssen,R.A.
  TITLE     Sorghum genome sequencing by methylation filtration
  JOURNAL   PLoS Biol. 3 (1), e13 (2005)
   PUBMED   15660154
COMMENT     Contact: Bedell JA
            Orion Genomics, LLC
            4041 Forest Park Ave, St. Louis, MO 63108, USA
            Tel: 314 615 6979
            Fax: 314 615 5975
            Email: jbedell@oriongenomics.com
            Plate: fsbb001f170  row: m  column: 16
            Seq primer: k Reverse
            Class: methylation filtered
            High quality sequence stop: 296.
FEATURES             Location/Qualifiers
     source          1. .296
                     /organism="Sorghum bicolor"
                     /mol_type="genomic DNA"
                     /cultivar="ATx623"
                     /db_xref="taxon:4558"
                     /clone="fsbb001f170m16"
                     /clone_lib="Sorghum methylation filtered library (LibID:
                     104)"
                     /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
                     prepared from purified nuclei was randomly sheared,
                     end-repaired, size fractionated to enrich for the 0.5 to 5
                     kb fraction, ligated into HincII-digested pBCSK(-) vector
                     and electroporated into E. coli cells. This is a
                     methylation filtered library."
ORIGIN

  Query Match            29.9%;  Score 47;  DB 10;  Length 296;
  Best Local Similarity  85.3%;  Pred. No. 0.00042;
  Matches   64;  Conservative    0;  Mismatches   10;  Indels    1;  Gaps    1;

Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT-CCATCTTACTCATGCA 122
              || || |||||||||||||||||| |||||||||||||||||| |||| |||| ||||||
Db        221 CCCTGGCTATTCTGAACCAAGAGGACACCTACTCCCAAACAATCCCATGTTACCCATGCA 280

Qy        123 ACTTCCATGCAAACA 137
              ||||||    ||||
Db        281 ACTTCCCATGCAACA 295


RESULT 6
CE588104/c
LOCUS       CE588104                636 bp    DNA     linear   GSS 28-SEP-2003
DEFINITION  tigr-gss-dog-17000366359543 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE588104
VERSION     CE588104.1  GI:36904885
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match            24.1%;  Score 37.8;  DB 10;  Length 636;
```

```
  Best Local Similarity   58.4%;  Pred. No. 0.48;
  Matches   66;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          45 TCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACA 104
               | || || |   || |||| | | | | ||  || ||||| | ||||||  |    || |
Db         292 TATTTTTTCCCAGTTCAAGGCATCCTTGTTTGGACCCAAGGGTATACCTCATTTTTAAAA 233

Qy         105 ATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
               || ||| ||| |   | || || ||  || | |   | |   || |  |||||
Db         232 ATACATTTTATTATAGAAATTTTCAAACACAAAAATAAAATGGTGTAGTGAAC 180


RESULT 7
CO689495/c
LOCUS       CO689495                643 bp    mRNA    linear   EST 26-JUL-2004
DEFINITION  DG11-25n6 DG11-kidney Canis familiaris cDNA 3', mRNA sequence.
ACCESSION   CO689495
VERSION     CO689495.1  GI:50638161
KEYWORDS    EST.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 643)
  AUTHORS   Schlueter,T., Hermanns,J., Weindel,M., Schuette,D., Kranz,H.,
            Henrich,J. and Loebbert,R.
  TITLE     Dog arrayTAG cDNA clone collection
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Thomas Schlueter
            LION bioscience AG
            Waldhoferstrasse 98, D-69123 Heidelberg, Germany
            Tel: +49 6221 4038 150
            Fax: +49 6221 4038 290
            Email: Thomas.Schlueter@lionbioscience.com.
FEATURES             Location/Qualifiers
     source          1. .643
                     /organism="Canis familiaris"
                     /mol_type="mRNA"
                     /strain="Beagle"
                     /db_xref="taxon:9615"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="DG11-kidney"
                     /note="Organ: kidney; Vector: Dog pBluescript LION"
ORIGIN

  Query Match             24.1%;  Score 37.8;  DB 7;  Length 643;
  Best Local Similarity   58.4%;  Pred. No. 0.49;
  Matches   66;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          45 TCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACA 104
               | || || |   || |||| | | | | ||  || ||||| | ||||||  |    || |
Db         333 TATTTTTTCCCAGTTCAAGGCATCCTTGTTTGGACCCAAGGGTATACCTCATTTTTAAAA 274

Qy         105 ATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
               || ||| ||| |   | || || ||  || | |   | |   || |  |||||
Db         273 ATACATTTTATTATAGAAATTTTCAAACACAAAAATAAAATGGTGTAGTGAAC 221


RESULT 8
AG366773/c
LOCUS       AG366773                1085 bp    DNA     linear   GSS 21-DEC-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-170B12.TJ, genomic survey
            sequence.
ACCESSION   AG366773
VERSION     AG366773.1  GI:47977978
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus (Japanese wild mouse)
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
            Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and
            Shiroishi,T.
  TITLE     Contribution of Asian mouse subspecies Mus musculus molossinus to
            genomic constitution of strain C57BL/6J, as defined by BAC-end
            sequence-SNP analysis
  JOURNAL   Genome Res. 14 (12), 2439-2447 (2004)
   PUBMED   15574823
REFERENCE   2  (bases 1 to 1085)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .1085
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-170B12.TJ"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             24.1%;  Score 37.8;  DB 10;  Length 1085;
  Best Local Similarity   51.9%;  Pred. No. 0.57;
  Matches   69;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy          22 CTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACC 81
               || |   | |  ||  ||   |  ||| | || || ||| | |||  ||| |||| ||
Db         719 CTCCCCTCCTCCTCCCCCTCCCTCCTTCTCCCCTCCTCCTACCCTCTCCTTTTCTATACT 660

Qy          82 AAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCA 141
               |     ||| |  | ||| |  || |||    |  |  |||| ||      ||  | |
Db         659 ACATNAATAACNCCCCCCNANNAAACCANNANAACCCCGCAATTTTTTNANAANNATAAA 600

Qy         142 CATATGTTTCCTG 154
               || |  |||  ||
Db         599 CAAACTTTTTTG 587


RESULT 9
CZ727957
LOCUS       CZ727957                874 bp    DNA     linear   GSS 25-JUL-2005
DEFINITION  OC__Ba0055L24.f OC__Ba Oryza coarctata genomic clone OC__Ba0055L24
            5', genomic survey sequence.
ACCESSION   CZ727957
VERSION     CZ727957.1  GI:71136431
KEYWORDS    GSS.
SOURCE      Oryza coarctata (Porteresia coarctata)
  ORGANISM  Oryza coarctata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
```

```
REFERENCE   1  (bases 1 to 874)
  AUTHORS   Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
  TITLE     OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0055  row: L  column: 24
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .874
                     /organism="Oryza coarctata"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:77588"
                     /clone="OC__Ba0055L24"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OC__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             22.7%;  Score 35.6;  DB 10;  Length 874;
  Best Local Similarity   78.6%;  Pred. No. 2.7;
  Matches   55;  Conservative    0;  Mismatches   14;  Indels    1;  Gaps    1;

Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 123
              ||||| |||||| ||||||| ||| |||||||| |||| |||   |  ||||||||||
Db        778 CCTTGGCTATTCCGAACCAACAGG-CACCTACTCTCAAAACATCACGCCGACTCATGCAA 836

Qy        124 CTTCCATGCA 133
              | ||||| |
Db        837 CCTCCATGTA 846


RESULT 10
BB505306
LOCUS       BB505306                700 bp    mRNA    linear   EST 25-OCT-2001
DEFINITION  BB505306 RIKEN full-length enriched, 10 days lactation, adult
            female mammary gland Mus musculus cDNA clone D730002G06, mRNA
            sequence.
ACCESSION   BB505306
VERSION     BB505306.2  GI:16442791
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     On Jul 27, 2000 this sequence version replaced gi:9514268.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
            Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="D730002G06"
                     /sex="female"
                     /tissue_type="mammary gland"
                     /dev_stage="10 days lactation, adult"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 10 days lactation,
                     adult female mammary gland"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                     was cleaved with BamHI and XhoI. Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC I."
ORIGIN

  Query Match             21.7%;  Score 34;  DB 2;  Length 700;
  Best Local Similarity   54.9%;  Pred. No. 8.3;
  Matches   67;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;

Qy         27 TGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAG 86
              | ||| |||  || | | |||  || | | || |    ||| | | ||        |||
Db        195 TTCTTTTTCTCCCCTCCCTCTGATTTCTTTGTACCTCACTTTCATTTTGGCCTTTAAGTT 254

Qy         87 GATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATAT 146
              | |   | |||||||| || |||| || | |       |   ||| | ||||| |  | ||
```

```
Db        255 GTTCTAGATTCCCAAAGAAGCCATTTTGCACTACAGTCAAACATACTAACACCCTGAAAT 314

Qy        147 GT 148
              ||
Db        315 GT 316


RESULT 11
CW871670/c
LOCUS       CW871670               734 bp    DNA     linear   GSS 12-FEB-2005
DEFINITION  she2h63-44.g_013.ab1 Whole-genome shotgun library of the elephant
            shark (aka elephant fish) Callorhinchus milii genomic, genomic
            survey sequence.
ACCESSION   CW871670
VERSION     CW871670.1  GI:59698305
KEYWORDS    GSS.
SOURCE      Callorhinchus milii (elephantfish)
  ORGANISM  Callorhinchus milii
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
            Holocephali; Chimaeriformes; Callorhinchidae; Callorhinchus.
REFERENCE   1  (bases 1 to 734)
  AUTHORS   Venkatesh,B., Tay,A., Dandona,N., Patil,J.G. and Brenner,S.
  TITLE     A compact cartilaginous fish model genome
  JOURNAL   Curr. Biol. 15 (3), R82-R83 (2005)
   PUBMED   15694293
COMMENT     Contact: Venkatesh B
            Molecular Genetics Lab
            Institute of Molecular and Cell Biology
            61 Biopolis Drive, Singapore 138673
            Tel: 65 6586 9571
            Fax: 65 6779 1117
            Email: mcbbv@imcb.a-star.edu.sg
            Whole-genome shotgun sequences of the elephant shark (aka elephant
            fish)
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .734
                     /organism="Callorhinchus milii"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7868"
                     /sex="Male"
                     /tissue_type="Testis"
                     /clone_lib="Whole-genome shotgun library of the elephant
                     shark (aka elephant fish)"
ORIGIN

  Query Match             21.5%;  Score 33.8;  DB 10;  Length 734;
  Best Local Similarity   54.4%;  Pred. No. 9.8;
  Matches   68;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              |||| ||| |||  | |||  | || || || |   | ||  | | |||||  ||   ||
Db        238 CGTGGCATTTCAATTTGCACTCCACGTGATTATGGCATCGAACTTGTTGTTAAATGACCC 179

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              | |    ||        ||      || || ||| |||| |||  |||| | || ||
Db        178 GAACTACACCACCCACAGACACCACAGACACATACACCCACACACACCATTTAACACAGA 119

Qy        121 CAACT 125
              | ||
Db        118 AATCT 114


RESULT 12
CL026582
LOCUS       CL026582               1677 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-23G24_Sp6.1 CH216 Xenopus tropicalis genomic clone
            CH216-23G24, genomic survey sequence.
ACCESSION   CL026582
VERSION     CL026582.1  GI:40470443
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1677)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 125
            High quality sequence stop: 217.
FEATURES             Location/Qualifiers
     source          1. .1677
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-23G24"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             21.4%;  Score 33.6;  DB 10;  Length 1677;
  Best Local Similarity   53.9%;  Pred. No. 15;
  Matches   69;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy          3 TGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAA 62
              |||| |||||||||   |    | || || ||   |  |||  | |||  |
Db        769 TGTCGTCTCACATGCACTTACTCCTGATTCTTTTGCTCTTCTACATGTGTCCATCCGGTT 828

Qy         63 GCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCA 122
              | |   ||||||| |      | |  |||| ||| ||     |  || ||| |||  ||
Db        829 TCATCAACTATTCTTACAGTTCATGTCACCTCCTCTCACTTTCTATATATTATTCAGACA 888

Qy        123 ACTTCCAT 130
              ||||||
Db        889 TCTTCCAT 896


RESULT 13
AI744861
LOCUS       AI744861               440 bp    mRNA    linear   EST 21-JUN-1999
DEFINITION  tr16b06.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2218451 3',
            mRNA sequence.
ACCESSION   AI744861
VERSION     AI744861.1  GI:5113149
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
```

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 411.

FEATURES Location/Qualifiers
source 1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2218451"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Ov23"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"

ORIGIN

Query Match 20.9%; Score 32.8; DB 1; Length 440;
Best Local Similarity 56.5%; Pred. No. 18;
Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```
Qy         46 CTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAA 105
              |||   ||| |   | ||  |  | |||||     ||||      | |||| || |   |
Db        259 CTTTGGCCAGCAAACTAGTATATCATATTCCTCCCCAAAGCAACACCTCCTACTTTATTT 318

Qy        106 TCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCT 153
              |||||||||| | |  || | || |||     |  || ||||||||
Db        319 TCCATCTTAATTAAACAGCGTCGATGTCTGCCTAGATATCTGTTTCCT 366
```

RESULT 14
CX446438/c
LOCUS CX446438 648 bp mRNA linear EST 07-JAN-2005
DEFINITION JGI_XZG22009.rev NIH_XGC_tropGas7 Xenopus tropicalis cDNA clone IMAGE:7537449 3', mRNA sequence.
ACCESSION CX446438
VERSION CX446438.1 GI:57262270
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 648)
AUTHORS Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C., Brokstein,P. and Lindquist,E.A.
TITLE DOE Joint Genome Institute Xenopus tropicalis EST project
JOURNAL Unpublished (2004)
COMMENT Other_ESTs: JGI_XZG22009.fwd
Contact: Lindquist,E.A., Richardson,P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Richard M. Harland Laboratory, University of California, Berkeley: http://tropicalis.berkeley.edu/home
cDNA Library Preparation: Richard M. Harland Laboratory, University of California, Berkeley
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL: http://image.llnl.gov
Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.rev' indicates a reverse sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
Poly-A: Based upon the presence of a run of 14 or more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed.
Plate: XZG 0229 row: b column: 7
High quality sequence stop: 185
POLYA=Yes.

FEATURES Location/Qualifiers
source 1. .648
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7537449"
/tissue_type="whole embryo"
/dev_stage="Gastrula (st. 10.5-12.5)"
/lab_host="E. coli XL1-Blue derivative, Stratagene ElectroTen-Blue"
/clone_lib="NIH_XGC_tropGas7"
/note="Vector: pCS108; Site_1: SalI; Site_2: NotI; Gastrula library constructed by Russell B. Fletcher in R. Harland's lab using poly A RNA and oligo dT primers (Invitrogen SuperScript Plasmid System for cDNA Synthesis and Cloning). SalI (5' end) -NotI (3' end) cDNA was inserted into vector pCS108 (http://mcb.berkeley.edu/labs/harland/pages/plasmids.html) ."

ORIGIN

Query Match 20.8%; Score 32.6; DB 8; Length 648;
Best Local Similarity 63.3%; Pred. No. 23;
Matches 50; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

```
Qy         74 TCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCA 133
              | ||||  | ||  |  ||  |||||| | | ||||| ||| || | || ||||||
Db        586 TTTGAATTATGAAAAGGCCATCTCCCATAGACTCCATTTTAGGCAAGTAATTTCCATTTT 527

Qy        134 AACACGCACATATGTTTCC 152
              || | | |  | |.|||||
Db        526 AAAAAGTATTTTTTTTTCC 508
```

RESULT 15
CW558661/c
LOCUS CW558661 740 bp DNA linear GSS 22-OCT-2004
DEFINITION OA_ABa0080B07.f OA_ABa Oryza australiensis genomic clone OA_ABa0080B07 5', genomic survey sequence.
ACCESSION CW558661
VERSION CW558661.1 GI:54424196
KEYWORDS GSS.
SOURCE Oryza australiensis
ORGANISM Oryza australiensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 740)
AUTHORS Kim,H., Yu,Y., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
TITLE OMAP Project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0080 row: B column: 07
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.

FEATURES Location/Qualifiers
source 1. .740
/organism="Oryza australiensis"
/mol_type="genomic DNA"

```
                     /db_xref="taxon:4532"
                     /clone="OA_ABa0080B07"
                     /tissue_type="young leaves"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="OA_ABa"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             20.8%;  Score 32.6;  DB 10;  Length 740;
  Best Local Similarity   63.3%;  Pred. No. 24;
  Matches   50;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          79 ACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACAC 138
               ||||| |    |  |||||||| |||||| ||||||  | |  ||| |    | ||||  |
Db         653 ACCAATATACTTACTACTCCCCAACAATGCATCTTCTTAAGCCAAATGGAAGGCAACAAT 594

Qy         139 GCACATATGTTTCCTGAAC 157
               |||  |  ||   |||| |
Db         593 GCAACTGCGTGAGCTGAGC 575


Search completed: March  6, 2006, 01:57:49
Job time : 742.207 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 21:54:12 ; Search time 43.8947 Seconds
(without alignments)
6357.883 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	157	100.0	1394	3	US-08-880-499-1	Sequence 1, Appli
	2	157	100.0	1394	3	US-08-880-499-2	Sequence 2, Appli
c	3	30.4	19.4	385136	3	US-09-949-016-16073	Sequence 16073, A
c	4	29.6	18.9	1185	3	US-09-543-681A-1081	Sequence 1081, Ap
c	5	29.2	18.6	1830121	3	US-09-557-884-1	Sequence 1, Appli
c	6	29.2	18.6	1830121	3	US-09-643-990A-1	Sequence 1, Appli
c	7	29.2	18.6	1830121	3	US-10-158-865-1	Sequence 1, Appli
c	8	28.8	18.3	64190	3	US-09-949-016-14712	Sequence 14712, A
c	9	28.8	18.3	64190	3	US-09-949-016-14713	Sequence 14713, A
	10	28	17.8	11022	3	US-09-949-016-15502	Sequence 15502, A
	11	28	17.8	41815	3	US-09-949-016-17447	Sequence 17447, A
c	12	27.8	17.7	601	3	US-09-949-016-191201	Sequence 191201,
c	13	27.8	17.7	11490	3	US-09-949-016-14149	Sequence 14149, A
	14	27.8	17.7	14952	3	US-09-949-016-16740	Sequence 16740, A
	15	27.6	17.6	927	3	US-09-937-862B-37	Sequence 37, Appl
c	16	27.4	17.5	1413	3	US-09-710-279-1813	Sequence 1813, Ap
c	17	27.4	17.5	1428	3	US-09-134-001C-1413	Sequence 1413, Ap
c	18	27.4	17.5	1794	3	US-09-270-767-10489	Sequence 10489, A
c	19	27.4	17.5	2472	2	US-08-425-299A-2	Sequence 2, Appli
	20	27.4	17.5	2926	3	US-09-710-279-3747	Sequence 3747, Ap
c	21	27.4	17.5	3014	3	US-09-710-279-3680	Sequence 3680, Ap
	22	27.4	17.5	3599	3	US-09-710-279-4003	Sequence 4003, Ap
c	23	27.2	17.3	3445	3	US-09-949-016-2556	Sequence 2556, Ap
c	24	27.2	17.3	3588	3	US-09-566-921-23	Sequence 23, Appl
c	25	27.2	17.3	6751	2	US-07-882-925A-5	Sequence 5, Appli
c	26	27.2	17.3	6751	2	US-08-184-012C-5	Sequence 5, Appli
c	27	27.2	17.3	8939	3	US-09-949-016-12703	Sequence 12703, A
c	28	27.2	17.3	9076	3	US-09-949-016-17368	Sequence 17368, A
	29	27.2	17.3	48994	3	US-09-949-016-14091	Sequence 14091, A
	30	27.2	17.3	60593	3	US-09-949-016-13779	Sequence 13779, A
c	31	27.2	17.3	62776	3	US-09-949-016-17576	Sequence 17576, A
c	32	27.2	17.3	100848	3	US-09-596-002-39	Sequence 39, Appl
c	33	27.2	17.3	158735	3	US-09-949-016-11989	Sequence 11989, A
c	34	27.2	17.3	158735	3	US-09-949-016-17130	Sequence 17130, A
c	35	27	17.2	879	3	US-09-134-001C-678	Sequence 678, App
	36	27	17.2	44180	3	US-09-949-016-12088	Sequence 12088, A
	37	27	17.2	47030	3	US-09-949-016-13037	Sequence 13037, A
	38	27	17.2	47030	3	US-09-949-016-13038	Sequence 13038, A
	39	27	17.2	47030	3	US-09-949-016-15039	Sequence 15039, A
	40	27	17.2	47030	3	US-09-949-016-15040	Sequence 15040, A
c	41	27	17.2	160552	3	US-09-593-828-11	Sequence 11, Appl
	42	27	17.2	194537	3	US-09-949-016-12928	Sequence 12928, A
	43	27	17.2	201529	3	US-09-949-016-12740	Sequence 12740, A
	44	26.8	17.1	265	3	US-09-513-999C-8379	Sequence 8379, Ap
c	45	26.8	17.1	271	3	US-08-946-026-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-880-499-1
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-880-499-1

Query Match 100.0%; Score 157; DB 3; Length 1394;

```
  Best Local Similarity   100.0%;  Pred. No. 9.4e-44;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 2
US-08-880-499-2
; Sequence 2, Application US/08880499
; Patent No. 6037523
;  GENERAL INFORMATION:
;    APPLICANT:  Albertson, Marc C.
;    APPLICANT:  Fox, Tim W.
;    APPLICANT:  Carl, Garnaat W.
;    APPLICANT:  Huffman, Gary A.
;    APPLICANT:  Kendall, Timmy L.
;    TITLE OF INVENTION:  MALE TISSUE-PREFERRED REGULATORY REGION
;    TITLE OF INVENTION:  AND METHOD OF USING SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PIONEER HI-BRED INTERNATIONAL, INC.
;      STREET:  Darwin Building, 7100 N.W. 62nd Ave., P.O.
;      STREET:  Box 1000
;      CITY:  Johnston
;      STATE:  Iowa
;      COUNTRY:  USA
;      ZIP:  50131
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/880,499
;      FILING DATE:  CONCURRENTLY HEREWITH
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sweeney, Patricia A.
;      REGISTRATION NUMBER:  32,733
;      REFERENCE/DOCKET NUMBER:  0578
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 248-4800
;      TELEFAX:  (515) 248-4844
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1394 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-880-499-2

  Query Match             100.0%;  Score 157;  DB 3;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 9.4e-44;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 3
US-09-949-016-16073/c
; Sequence 16073, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16073
;   LENGTH: 385136
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(385136)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16073

  Query Match             19.4%;  Score 30.4;  DB 3;  Length 385136;
  Best Local Similarity   57.3%;  Pred. No. 14;
  Matches   55;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy         55 TCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTA 114
              || ||||   |   |||||| ||||  |     |  | ||| | ||||||  |     ||
Db     253908 TCCTCCAGATCGCTCCTATTTTGAAATATATAAACGCTTACACACAAACACACTCATATA 253849

Qy        115 CTCATGCAACTTCCATGCAAACACGCACATATGTTT 150
                |||   ||| |||    | ||| ||||| | |||
Db     253848 TCCATTTTACTCCCACCACACCACACACATTTTTTT 253813


RESULT 4
US-09-543-681A-1081/c
; Sequence 1081, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1081
;   LENGTH: 1185
;   TYPE: DNA
;   ORGANISM: Proteus mirabilis
US-09-543-681A-1081

  Query Match             18.9%;  Score 29.6;  DB 3;  Length 1185;
  Best Local Similarity   59.5%;  Pred. No. 3.4;
  Matches   50;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy         35 CAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTA 94
```

```
              |||| || || ||  |||| |    | ||| || ||| |  |||||   ||| |||
Db        940 CAACAATTTGTTTTTGTCCACCAGAAAGACCTAACCCATTTTCTCCCAAGGACATATCTA 881

Qy         95 CTCCCAAACAATCCATCTTACTCA 118
              ||||  |  || | | ||| | |
Db        880 ATCCCCTAGGATGCTTTTTAATAA 857


RESULT 5
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
;  GENERAL INFORMATION:
;       APPLICANT: Fleischmann et al.
;       TITLE OF INVENTION: The Nucleotide sequence of
;                           the Haemophilus influenzae Rd Genome, Fragments
;                           Thereof, and Uses Thereof
;       NUMBER OF SEQUENCES: 1
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: MD
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3 1/2 inch diskette
;            COMPUTER: Dell Pentium
;            OPERATING SYSTEM: MS DOS v6.22
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,884
;            FILING DATE: 25-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/476,102
;            FILING DATE: JUN-5-1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Michelle S. Marks
;            REGISTRATION NUMBER: 41,971
;            REFERENCE/DOCKET NUMBER: PB186P3
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 301-309-8504
;            TELEFAX: 301-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1830121 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

  Query Match           18.6%;  Score 29.2;  DB 3;  Length 1830121;
  Best Local Similarity  51.5%;  Pred. No. 63;
  Matches   67;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy         12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
              || |   || |  | ||||||  |    | |  ||         | |||||  | ||||
Db     677275 ACTTTAAATTCAGCCTGCTTGCTTGCTCTATTGCATTTGCATTAAATGCAAGCATAGCCT 677216

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              || |||  | |  |   ||| ||   | ||| | ||| |  ||  |   | |  |||
Db     677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156

Qy        132 CAAACACGCA 141
              |||.||  ||
Db     677155 CAACCAACCA 677146


RESULT 6
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
;  GENERAL INFORMATION:
;       APPLICANT: Robert D. Fleischmann
;                  Mark D. Adams
;                  Owen White
;                  Hamilton O. Smith
;                  J. Craig Venter
;       TITLE OF INVENTION: The Nucleotide sequence of
;                           the Haemophilus influenzae Rd Genome, Fragments
;                           Thereof, and Uses Thereof
;       NUMBER OF SEQUENCES: 1
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville,
;            STATE: MD
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3 1/2 inch diskette
;            COMPUTER: Dell Pentium
;            OPERATING SYSTEM: MS DOS v6.22
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/643,990A
;            FILING DATE: 23-Aug-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/487,429
;            FILING DATE: 1995-06-07
;            APPLICATION NUMBER: 08/426,787
;            FILING DATE: 1995-04-21
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Kenley K. Hoover
;            REGISTRATION NUMBER: 40,302
;            REFERENCE/DOCKET NUMBER: PB186P1C1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 301-610-5790
;            TELEFAX: 310-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1830121 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

  Query Match           18.6%;  Score 29.2;  DB 3;  Length 1830121;
  Best Local Similarity  51.5%;  Pred. No. 63;
  Matches   67;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy         12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
              || |   || |  | ||||||  |    | |  ||         | |||||  | ||||
Db     677275 ACTTTAAATTCAGCCTGCTTGCTTGCTCTATTGCATTTGCATTAAATGCAAGCATAGCCT 677216

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              || |||  | |  |   ||| ||   | ||| | ||| |  ||  |   | |  |||
Db     677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156

Qy        132 CAAACACGCA 141
              |||.||  ||
Db     677155 CAACCAACCA 677146


RESULT 7
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fragm
; Patent No. 6846651
```

; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature

```
;   LOCATION: (122336)..(122336)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131340)..(131340)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131360)..(131360)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (139910)..(139910)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (140398)..(140398)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (142750)..(142750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145058)..(145058)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145171)..(145171)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145942)..(145942)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (147197)..(147197)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (150841)..(150841)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152500)..(152500)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature

  Query Match             18.6%;  Score 29.2;  DB 3;  Length 1830121;
  Best Local Similarity   51.5%;  Pred. No. 63;
  Matches   67;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy          12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
               || |   || | | ||||||| |     | | | ||          |||||| | ||||
Db      677275 ACTTTAAATTCAGCCTGCTTGCTTGCTCTATTGCATTTGCATTAAATGCAAGCATAGCCT 677216

Qy          72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
               || |||  | | |    ||| ||   | ||| | ||| |  ||  |   | |  |||
Db      677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156

Qy         132 CAAACACGCA 141
               ||| ||  ||
Db      677155 CAACCAACCA 677146


RESULT 8
US-09-949-016-14712/c
; Sequence 14712, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14712
;   LENGTH: 64190
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14712

  Query Match             18.3%;  Score 28.8;  DB 3;  Length 64190;
  Best Local Similarity   62.5%;  Pred. No. 27;
  Matches   45;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy          55 TCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTA 114
               || |||||   || ||  | || | ||| || |  ||  | | | || ||| |||   |
Db         355 TCATCCAATAGTTTCCACTACTAATCCATGATGGAACTGAGTACTAATCAAGCCAATATT 296

Qy         115 CTCATGCAACTT 126
               |||||| ||| |
Db         295 CTCATGTAACAT 284


RESULT 9
US-09-949-016-14713/c
; Sequence 14713, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14713
;   LENGTH: 64190
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14713

  Query Match             18.3%;  Score 28.8;  DB 3;  Length 64190;
  Best Local Similarity   62.5%;  Pred. No. 27;
  Matches   45;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy          55 TCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTA 114
               || |||||   || ||  | || | ||| || |  ||  | | | || ||| |||   |
Db         355 TCATCCAATAGTTTCCACTACTAATCCATGATGGAACTGAGTACTAATCAAGCCAATATT 296

Qy         115 CTCATGCAACTT 126
               |||||| ||| |
Db         295 CTCATGTAACAT 284


RESULT 10
US-09-949-016-15502
; Sequence 15502, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15502
;   LENGTH: 11022
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15502

  Query Match             17.8%;  Score 28;  DB 3;  Length 11022;
  Best Local Similarity   63.2%;  Pred. No. 27;
  Matches   43;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         15 TGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATT 74
              |||| || | |||| |  ||  |  |||  | ||||  ||||      ||||||| | |
Db       4432 TGGCTTATTTCATGGTGTTTTTAGGGTTAATGTTGTAGCATCCATTTAGCCTTGTCAAGG 4491

Qy         75 CTGAACCA 82
              ||||| ||
Db       4492 CTGAATCA 4499


RESULT 11
US-09-949-016-17447
; Sequence 17447, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17447
;   LENGTH: 41815
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17447

  Query Match             17.8%;  Score 28;  DB 3;  Length 41815;
  Best Local Similarity   58.3%;  Pred. No. 43;
  Matches   49;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy         62 AGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGC 121
              || | |||  |||| |  ||   | ||| || |||| |   |  ||   | ||||||| |
Db      33058 AGACATGCAGATTCACACACATCCGCATAGCTGCTCCTACGGATCCCTGGTCACTCATTC 33117

Qy        122 AACTTCCATGCAAACACGCACATA 145
              |  |  |||  | ||| |||  ||
Db      33118 ATGTCACATAGAGACATGCATGTA 33141


RESULT 12
```

```
US-09-949-016-191201/c
; Sequence 191201, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 191201
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-191201

  Query Match             17.7%;  Score 27.8;  DB 3;  Length 601;
  Best Local Similarity   58.0%;  Pred. No. 11;
  Matches   47;  Conservative    1;  Mismatches   33;  Indels    0;  Gaps    0;

Qy         57 GTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACT 116
              ||||   | ||||  |   |  :  || | | |   || | || ||| ||  ||| | |
Db        322 GTCCTGACGTTGCATGCACACRTGCACGCGCACAAACACACACACACACTCTCTCTCATT 263

Qy        117 CATGCAACTTCCATGCAAACA 137
              ||||||| |   ||| ||| |
Db        262 CATGCAAATAGGATGGAAATA 242


RESULT 13
US-09-949-016-14149/c
; Sequence 14149, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14149
;   LENGTH: 11490
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14149

  Query Match             17.7%;  Score 27.8;  DB 3;  Length 11490;
  Best Local Similarity   52.1%;  Pred. No. 32;
  Matches   62;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy         38 CCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTC 97
              | |  | |||  |||| |   ||   || ||||    |   |||| |||||  || || |
Db       5642 CAGGGCCTCTGCTTCCCTGCCCCTCCCCCTGCCCCCACCTCACCAGGAGGAAGCCCACGC 5583

Qy         98 CCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAA 156
              ||   ||  ||   |  | ||| |   |      |||  |  |   | ||||| ||  ||
Db       5582 CCTTGCACACCTGATGCCACATCCGCTTGAGTGGCAGGCGAGTGAAGATGTTGCCCAAA 5524
```

```
RESULT 14
US-09-949-016-16740
; Sequence 16740, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16740
;   LENGTH: 14952
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16740

  Query Match           17.7%;  Score 27.8;  DB 3;  Length 14952;
  Best Local Similarity   59.5%;  Pred. No. 35;
  Matches   47;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy         67 TGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTT 126
              || | |   ||| |||| || |||  || | |  |  ||| | |||  | || | ||||
Db         70 TGTCAACCGTGATCCAACAGTATAATTAGTTCTTAGGAATTCCTCTCTCACAGGTAACTA 129

Qy        127 CCATGCAAACACGCACATA 145
                |  | |||| |||| |
Db        130 GGTTTTACACACACACACA 148


RESULT 15
US-09-937-862B-37
; Sequence 37, Application US/09937862B
; Patent No. 6846621
; GENERAL INFORMATION:
;  APPLICANT: Oberste, M. Steven
;  APPLICANT:  Maher, Kaija
;  APPLICANT:  Kilpatrick, David R.
;  APPLICANT:  Pallansch, Mark A.
;  TITLE OF INVENTION: TYPING OF HUMAN ENTEROVIRUSES
;  FILE REFERENCE: 14114.0353U2
;  CURRENT APPLICATION NUMBER: US/09/937,862B
;  CURRENT FILING DATE:  2001-09-28
;  PRIOR APPLICATION NUMBER: PCT/US00/07828
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/127,464
;  PRIOR FILING DATE: 1999-03-31
;  NUMBER OF SEQ ID NOS: 89
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 37
;   LENGTH: 927
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence; No. 6846621e =
;   OTHER INFORMATION: synthetic construct
US-09-937-862B-37

  Query Match           17.6%;  Score 27.6;  DB 3;  Length 927;
  Best Local Similarity   55.1%;  Pred. No. 15;
  Matches   54;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy         53 CATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCT 112
              |||| || || || |   ||||||  ||| | | | | | |||| |   || |||
Db        551 CATCATCAAACCCCTCAATATTCTACACCTATGGAACAGCACCACCCAGAATTTCGATCC 610

Qy        113 TACTCATGCAACTTCCATGCAAACACGCACATATGTTT 150
                 |  ||    || ||  |    || |||| | | |
Db        611 CTTTTGTGGGCATTACAAACGCGTACTCACATTTTTAT 648


Search completed: March  5, 2006, 22:36:28
Job time : 46.8947 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 23:29:31 ; Search time 225.751 Seconds
(without alignments)
5750.985 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	157	100.0	255	8	US-10-713-381-9	Sequence 9, Appli
2	157	100.0	1394	8	US-10-713-381-1	Sequence 1, Appli
3	157	100.0	1394	8	US-10-713-381-2	Sequence 2, Appli
4	146	93.0	158	8	US-10-713-381-3	Sequence 3, Appli
5	50	31.8	50	8	US-10-713-381-5	Sequence 5, Appli
6	40	25.5	40	8	US-10-713-381-6	Sequence 6, Appli
c 7	32.2	20.5	399	8	US-10-674-124A-9081	Sequence 9081, Ap
c 8	31.6	20.1	2352	7	US-10-282-122A-36272	Sequence 36272, A
9	31.4	20.0	565	4	US-09-925-065A-871704	Sequence 871704,
c 10	30.8	19.6	601	4	US-09-925-065A-778732	Sequence 778732,
c 11	30.8	19.6	1122	7	US-10-437-963-1496	Sequence 1496, Ap
12	30.6	19.5	559	5	US-10-027-632-244811	Sequence 244811,
13	30.6	19.5	559	6	US-10-027-632-244811	Sequence 244811,
14	30.6	19.5	79467	7	US-10-052-482-223	Sequence 223, App
15	30.4	19.4	52001	7	US-10-317-273-11	Sequence 11, Appl
16	30.2	19.2	436	4	US-09-925-065A-153052	Sequence 153052,
17	30.2	19.2	1266	9	US-10-501-282-2923	Sequence 2923, Ap
18	30.2	19.2	1754382	9	US-10-501-282-6651	Sequence 6651, Ap
19	30	19.1	30	8	US-10-713-381-4	Sequence 4, Appli
20	30	19.1	843	5	US-10-027-632-170626	Sequence 170626,
21	30	19.1	843	6	US-10-027-632-170626	Sequence 170626,
22	29.8	19.0	236	3	US-09-783-590-10490	Sequence 10490, A
23	29.8	19.0	617	4	US-09-925-065A-285326	Sequence 285326,
24	29.8	19.0	617	4	US-09-925-065A-285327	Sequence 285327,
25	29.8	19.0	617	4	US-09-925-065A-285328	Sequence 285328,
c 26	29.8	19.0	691	7	US-10-424-599-57018	Sequence 57018, A
27	29.8	19.0	1363	4	US-09-925-065A-12385	Sequence 12385, A
28	29.8	19.0	2735	5	US-10-027-632-254612	Sequence 254612,
29	29.8	19.0	2735	6	US-10-027-632-254612	Sequence 254612,
30	29.8	19.0	1223197	5	US-10-027-632-179264	Sequence 179264,
31	29.8	19.0	1223197	6	US-10-027-632-179264	Sequence 179264,
c 32	29.6	18.9	680	7	US-10-767-701-2154	Sequence 2154, Ap
c 33	29.2	18.6	1475	6	US-10-321-802-11	Sequence 11, Appl
c 34	29.2	18.6	1830121	7	US-10-329-670-1	Sequence 1, Appli
c 35	29.2	18.6	1830121	8	US-10-158-865-1	Sequence 1, Appli
c 36	29.2	18.6	1830121	9	US-10-981-687-1	Sequence 1, Appli
c 37	29	18.5	358	8	US-10-425-115-78606	Sequence 78606, A
38	29	18.5	553	5	US-10-027-632-222444	Sequence 222444,
39	29	18.5	553	5	US-10-027-632-222445	Sequence 222445,
40	29	18.5	553	6	US-10-027-632-222444	Sequence 222444,
41	29	18.5	553	6	US-10-027-632-222445	Sequence 222445,
c 42	29	18.5	7192	7	US-10-437-963-85365	Sequence 85365, A
43	28.8	18.3	544	4	US-09-925-065A-805392	Sequence 805392,
44	28.8	18.3	732	5	US-10-027-632-21721	Sequence 21721, A
45	28.8	18.3	732	5	US-10-027-632-21722	Sequence 21722, A

ALIGNMENTS

RESULT 1
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%; Score 157; DB 8; Length 255;
Best Local Similarity 100.0%; Pred. No. 2.1e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 74

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 134

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db        135 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171
```

RESULT 2
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1

```
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1394
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-1

  Query Match             100.0%;  Score 157;  DB 8;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-43;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 3
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1394
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-2

  Query Match             100.0%;  Score 157;  DB 8;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-43;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 4
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 158
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-3

  Query Match             93.0%;  Score 146;  DB 8;  Length 158;
  Best Local Similarity   99.4%;  Pred. No. 1e-39;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 59
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy         60 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120

Qy        120 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158


RESULT 5
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-5
```

```
  Query Match             31.8%;  Score 50;  DB 8;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-07;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         75 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 50


RESULT 6
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;    LENGTH: 40
;    TYPE: DNA
;    ORGANISM: Zea mays
US-10-713-381-6

  Query Match             25.5%;  Score 40;  DB 8;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         85 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||||||||||||||||||||||||||||||||||||||
Db          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40


RESULT 7
US-10-674-124A-9081/c
; Sequence 9081, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
;  APPLICANT: INOKO, Hidetoshi
;  APPLICANT:  TAMIYA, Gen
;  TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
;  TITLE OF INVENTION:  GENETIC POLYMORPHISM MARKERS
;  FILE REFERENCE: ORIN-003CIP
;  CURRENT APPLICATION NUMBER: US/10/674,124A
;  CURRENT FILING DATE:  2003-09-26
;  PRIOR APPLICATION NUMBER: 10/257,511
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: PCT/JP00/07621
;  PRIOR FILING DATE: 2000-10-30
;  PRIOR APPLICATION NUMBER: JP2000-112699
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: JP2002-327516
;  PRIOR FILING DATE: 2002-09-28
;  PRIOR APPLICATION NUMBER: JP2002-383869
;  PRIOR FILING DATE: 2002-12-09
;  NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 9081
;    LENGTH: 399
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    OTHER INFORMATION: chr5.fa.O7frz.139538206
;    FEATURE:
;    OTHER INFORMATION: Located on chromosome 5
;    FEATURE:
;    OTHER INFORMATION: Distance between a terminus base of telomere on
;    OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
;    OTHER INFORMATION: sequence : 128437212
;    FEATURE:
;    OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
;    OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
;    OTHER INFORMATION: 5'-terminus of this base sequence : 250186
US-10-674-124A-9081

  Query Match             20.5%;  Score 32.2;  DB 8;  Length 399;
  Best Local Similarity   61.2%;  Pred. No. 1.6;
  Matches   52;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              ||| |||| ||   |    | || | |||| || ||| |||||||  |||||  |
Db        239 ATTTTGAAGAAAACAGTGGGCCACCCTCAAAAAACCCAAACTACTCATCTAACTTAAAAA 180

Qy        132 CAAACACGCACATATGTTTCCTGAA 156
              ||| |  |||| || | |  ||||
Db        179 AAAAAAACCACACATTTGTGATGAA 155


RESULT 8
US-10-282-122A-36272/c
; Sequence 36272, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36272
;    LENGTH: 2352
;    TYPE: DNA
;    ORGANISM: Streptococcus mutans
```

US-10-282-122A-36272

Query Match 20.1%; Score 31.6; DB 7; Length 2352;
Best Local Similarity 58.5%; Pred. No. 4.9;
Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
Qy         62 AGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGC 121
              | |||  |      |  || | | ||||  | ||||||  || ||||  |  |||  |||
Db        234 ATCCTGACTGTGCCCAAAGCCAAAGGAAGCATACTCCTGAAAAATCTGTGCTACATCTGC 175

Qy        122 AACTTCCATGCAAACACGCACATATGTTTCCTGA 155
              || |||    ||  |||  ||  | ||||||||
Db        174 AATTTCAGCAAAAGAACGTTCAATTTTTTCCTGA 141
```

RESULT 9
US-09-925-065A-871704
; Sequence 871704, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 871704
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-871704

Query Match 20.0%; Score 31.4; DB 4; Length 565;
Best Local Similarity 60.2%; Pred. No. 3.4;
Matches 50; Conservative 1; Mismatches 32; Indels 0; Gaps 0;

```
Qy          4 GTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAG 63
              | | |||  | ||| | |   |||||||| |  | ||  | |||   ||| || |  ||
Db        274 GGCCTCTTGCTTGGAAGAGCCCATGCTTGGTTTAACGCCCTTCTGTCTCCGTCTTGAAAT 333

Qy         64 CCTTGCCTATTCTGAACCAAGAG 86
                | ||   ||| ||||| |:|||
Db        334 TCATGATAATTTTGAACAARGAG 356
```

RESULT 10
US-09-925-065A-778732/c
; Sequence 778732, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 778732
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-778732

Query Match 19.6%; Score 30.8; DB 4; Length 601;
Best Local Similarity 52.3%; Pred. No. 5.5;
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
Qy         12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
              |||| | | |||  ||   |  |  || ||| ||| | |   ||  |   | |||| |||
Db        354 ACATAGTACACTTTATTTATTGTTTACTGTTGGTCATTTAGGATATTTATACCCTTCCCT 295

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              ||| | ||  |  |   |  ||| | || ||| |   |||  |      | ||| | |
Db        294 ATTATAAATGATAAATTTTTCTATACACATACAGTTTTTCTACCCATATTATCTTTCCTA 235

Qy        132 CAAACACGCA 141
                | ||   ||
Db        234 AACACTGACA 225
```

RESULT 11
US-10-437-963-1496/c
; Sequence 1496, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 1496
; LENGTH: 1122
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101357C.1
US-10-437-963-1496

Query Match 19.6%; Score 30.8; DB 7; Length 1122;
Best Local Similarity 57.1%; Pred. No. 7;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy         60 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 119
              ||||||| || |  | || | |  | ||||  |||  | |||  || |||| ||  |  |
Db        313 CAAGCCTCGCATCATTTGCATCCCGGGGATCACTAAACTCAACAAAGCCATATTCATTCT 254

Qy        120 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
                 ||  | |         ||| | |  || | | ||||||
Db        253 TCAGATCCACAAATCGCACCCTCCCATATCTGCTGAAC 216
```

RESULT 12
US-10-027-632-244811
; Sequence 244811, Application US/10027632

```
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244811
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-244811

  Query Match             19.5%;  Score 30.6;  DB 5;  Length 559;
  Best Local Similarity   51.9%;  Pred. No. 6.3;
  Matches   69;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy         24 ACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAA 83
              ||||||| | |  | |  ||||  ||    |||| | |  |  | | | ||||  |  |
Db        380 ACATGCATTTCTAGCTTTTCGATTTTGCTCATCTTGCTGGAGGTACTTTTTCTAGATTCA 439

Qy         84 GAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACA 143
                | | |||||    ||| | |||||   | || | |||  |||       ||||  |
Db        440 TCTGCTCCCTACCTTTAAAGATTCCATGGAAGTCCTACAAAATCCCATAGCACACTGAAC 499

Qy        144 TATGTTTCCTGAA 156
                | | | |  ||
Db        500 ATTTTCTTCCCAA 512


RESULT 13
US-10-027-632-244811
; Sequence 244811, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244811
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-244811

  Query Match             19.5%;  Score 30.6;  DB 6;  Length 559;
  Best Local Similarity   51.9%;  Pred. No. 6.3;
  Matches   69;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy         24 ACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAA 83
              ||||||| | |  | |  ||||  ||    |||| | |  |  | | | ||||  |  |
Db        380 ACATGCATTTCTAGCTTTTCGATTTTGCTCATCTTGCTGGAGGTACTTTTTCTAGATTCA 439

Qy         84 GAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACA 143
                | | |||||    ||| | |||||   | || | |||  |||       ||||  |
Db        440 TCTGCTCCCTACCTTTAAAGATTCCATGGAAGTCCTACAAAATCCCATAGCACACTGAAC 499

Qy        144 TATGTTTCCTGAA 156
                | | | |  ||
Db        500 ATTTTCTTCCCAA 512


RESULT 14
US-10-052-482-223
; Sequence 223, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
;  APPLICANT: Engelhard, Eric
;  APPLICANT:  Morris, David
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
;  FILE REFERENCE: A-71087/RMS/DCF
;  CURRENT APPLICATION NUMBER: US/10/052,482
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 223
;   LENGTH: 79467
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4099)..(4369)
;   OTHER INFORMATION: "n" at positions 4099 to 4369 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5502)..(5521)
;   OTHER INFORMATION: "n" at positions 5502 to 5521 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10089)..(10620)
;   OTHER INFORMATION: "n" at positions 10089 to 10620 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (13273)..(13370)
;   OTHER INFORMATION: "n" at positions 13273 to 13370 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (20762)..(20781)
;   OTHER INFORMATION: "n" at positions 20762 to 20781 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (30751)..(30916)
;   OTHER INFORMATION: "n" at positions 30751 to 30916 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (46579)..(46772)
;   OTHER INFORMATION: "n" at positoins 46759 to 46772 can be any base
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (77382)..(77401)
;   OTHER INFORMATION: "n" at positions 77372 to 77401 can be any base
US-10-052-482-223

  Query Match             19.5%;  Score 30.6;  DB 7;  Length 79467;
  Best Local Similarity   56.4%;  Pred. No. 39;
  Matches   57;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy         51 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 110
              ||||    |  | ||||||  |  |||| | ||||  ||  | || || || |||  |
Db      53779 TCCACTAACTGATCCTTGCTTCCTCTGGAACAAGCTGAGTCTTAGTCACACACACACACA 53838

Qy        111 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTC 151
              |  || ||  ||    |||  || ||||   | | ||  ||
Db      53839 CACACACACACACACACCACACACACACAAGCTTGTGGCTC 53879


RESULT 15
US-10-317-273-11
; Sequence 11, Application US/10317273
; Publication No. US20040110158A1
; GENERAL INFORMATION:
;  APPLICANT: Kenneth W. Dobie
;  APPLICANT:  Tamara Balac Sipes
;  TITLE OF INVENTION: MODULATION OF MITOCHONDRIAL RIBOSOMAL PROTEIN L13 EXPRESSION
;  FILE REFERENCE: RTS-0478
;  CURRENT APPLICATION NUMBER: US/10/317,273
;  CURRENT FILING DATE:  2002-12-10
;  NUMBER OF SEQ ID NOS: 159
; SEQ ID NO 11
;   LENGTH: 52001
;   TYPE: DNA
;   ORGANISM: H. sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11121, 11122, 11123, 11124, 11125, 11126, 11127, 11128, 11129,
;   LOCATION: 11130, 11131, 11132, 11133, 11134, 11135, 11136, 11137,
;   LOCATION: 11138, 11139, 11140, 11141, 11142, 11143, 11144, 11145,
;   LOCATION: 11146, 11147, 11148, 11149, 11150, 11151, 11152, 11153
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11154, 11155, 11156, 11157, 11158, 11159, 11160, 11161, 11162,
;   LOCATION: 11163, 11164, 11165, 11166, 11167, 11168, 11169, 11170,
;   LOCATION: 11171, 11172, 11173, 11174, 11175, 11176, 11177, 11178,
;   LOCATION: 11179, 11180, 11181, 11182, 11183, 11184, 11185, 11186
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11187, 11188, 11189, 11190, 11191, 11192, 11193, 11194, 11195,
;   LOCATION: 11196, 11197, 11198, 11199, 11200, 11201, 11202, 11203,
;   LOCATION: 11204, 11205, 11206, 11207, 11208, 11209, 11210, 11211,
;   LOCATION: 11212, 11213, 11214, 11215, 11216, 11217, 11218, 11219
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11220
;   OTHER INFORMATION: n = A,T,C or G
US-10-317-273-11

  Query Match             19.4%;  Score 30.4;  DB 7;  Length 52001;
  Best Local Similarity   61.2%;  Pred. No. 39;
  Matches   49;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         50 TTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCA 109
              || | |  || ||  ||    |    ||||  ||||| ||| | | |   |||| | |||
Db      20055 TTTCTTTTTCTAAAACTGAAATTCCATGAAAAAAGAGAATATCAATTTTAAAACTACCCA 20114

Qy        110 TCTTACTCATGCAACTTCCA 129
              |||| ||||   |||||| |
Db      20115 TCTTTCTCAGCAAACTTCTA 20134
```

Search completed: March 6, 2006, 03:55:15
Job time : 226.751 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 01:58:03 ; Search time 163.89 Seconds
(without alignments)
2100.307 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	31.4	20.0	565	6	US-09-925-065A-871704	Sequence 871704,
c	2	30.8	19.6	601	6	US-09-925-065A-778732	Sequence 778732,
	3	30.6	19.5	64415	12	US-11-117-187-185	Sequence 185, App
c	4	30.4	19.4	191797	12	US-11-121-086-13	Sequence 13, Appl
	5	30.2	19.2	436	6	US-09-925-065A-153052	Sequence 153052,
	6	29.8	19.0	617	6	US-09-925-065A-285326	Sequence 285326,
	7	29.8	19.0	617	6	US-09-925-065A-285327	Sequence 285327,
	8	29.8	19.0	617	6	US-09-925-065A-285328	Sequence 285328,
	9	29.8	19.0	1363	6	US-09-925-065A-12385	Sequence 12385, A
c	10	29.4	18.7	611587	12	US-11-117-187-209	Sequence 209, App
	11	28.8	18.3	544	6	US-09-925-065A-805392	Sequence 805392,
	12	28.8	18.3	1370	8	US-10-750-185-47517	Sequence 47517, A
	13	28.8	18.3	1370	8	US-10-750-623-47517	Sequence 47517, A
	14	28.4	18.1	628	6	US-09-925-065A-845392	Sequence 845392,
c	15	28.4	18.1	1691	6	US-09-925-065A-33976	Sequence 33976, A
c	16	28.4	18.1	1691	6	US-09-925-065A-33977	Sequence 33977, A
c	17	28.4	18.1	1691	6	US-09-925-065A-33980	Sequence 33980, A
	18	28.2	18.0	2069	6	US-09-925-065A-724023	Sequence 724023,
	19	28	17.8	451	6	US-09-925-065A-173193	Sequence 173193,
c	20	28	17.8	801	6	US-09-925-065A-32050	Sequence 32050, A
c	21	28	17.8	801	6	US-09-925-065A-32051	Sequence 32051, A
c	22	28	17.8	1691	6	US-09-925-065A-33978	Sequence 33978, A
c	23	28	17.8	1691	6	US-09-925-065A-33979	Sequence 33979, A
c	24	28	17.8	1740	6	US-09-925-065A-553735	Sequence 553735,
c	25	28	17.8	1740	6	US-09-925-065A-553736	Sequence 553736,
c	26	28	17.8	2327	6	US-09-925-065A-686351	Sequence 686351,
c	27	28	17.8	2327	6	US-09-925-065A-686352	Sequence 686352,
c	28	28	17.8	2327	6	US-09-925-065A-686353	Sequence 686353,
c	29	28	17.8	2327	6	US-09-925-065A-686354	Sequence 686354,
c	30	28	17.8	2341	8	US-10-750-185-30272	Sequence 30272, A
c	31	28	17.8	2341	8	US-10-750-623-30272	Sequence 30272, A
	32	28	17.8	2750	8	US-10-750-185-39884	Sequence 39884, A
	33	28	17.8	2750	8	US-10-750-623-39884	Sequence 39884, A
	34	27.8	17.7	541	6	US-09-925-065A-427091	Sequence 427091,
c	35	27.8	17.7	560	6	US-09-925-065A-945788	Sequence 945788,
	36	27.8	17.7	576	6	US-09-925-065A-757513	Sequence 757513,
	37	27.8	17.7	1476	8	US-10-750-185-37809	Sequence 37809, A
	38	27.8	17.7	1476	8	US-10-750-623-37809	Sequence 37809, A
	39	27.8	17.7	3242	8	US-10-750-185-58624	Sequence 58624, A
	40	27.8	17.7	3242	8	US-10-750-623-58624	Sequence 58624, A
c	41	27.8	17.7	393323	7	US-10-330-773-23	Sequence 23, Appl
c	42	27.6	17.6	1724	8	US-10-750-185-55178	Sequence 55178, A
c	43	27.6	17.6	1724	8	US-10-750-623-55178	Sequence 55178, A
	44	27.6	17.6	2246	6	US-09-925-065A-68127	Sequence 68127, A
	45	27.4	17.5	531	6	US-09-925-065A-95915	Sequence 95915, A

ALIGNMENTS

RESULT 1
US-09-925-065A-871704
; Sequence 871704, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 871704
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-871704

Query Match 20.0%; Score 31.4; DB 6; Length 565;
Best Local Similarity 60.2%; Pred. No. 2.7;
Matches 50; Conservative 1; Mismatches 32; Indels 0; Gaps 0;

```
Qy          4 GTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAG 63
              | | ||| . | ||| | |   |||||||| |  | ||  | |||   ||| || |  ||
Db        274 GGCCTCTTGCTTGGAAGAGCCCATGCTTGGTTTAACGCCCTTCTGTCTCCGTCTTGAAAT 333

Qy         64 CCTTGCCTATTCTGAACCAAGAG 86
                | ||   ||| ||||| |:|||
Db        334 TCATGATAATTTTGAACAARGAG 356
```

RESULT 2

US-09-925-065A-778732/c
; Sequence 778732, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 778732
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-778732

Query Match 19.6%; Score 30.8; DB 6; Length 601;
Best Local Similarity 52.3%; Pred. No. 4.3;
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
Qy         12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
              |||| | | ||| || |  | | || ||| ||| | |  || |  | |||| |||
Db        354 ACATAGTACACTTTATTTATTGTTTACTGTTGGTCATTTAGGATATTTATACCCTTCCCT 295

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              ||| | ||  |  |   |  ||| | || ||| |  |||  |     | ||| | |
Db        294 ATTATAAATGATAAATTTTTCTATACACATACAGTTTTTCTACCCATATTATCTTTCCTA 235

Qy        132 CAAACACGCA 141
                | ||  ||
Db        234 AACACTGACA 225
```

RESULT 3
US-11-117-187-185
; Sequence 185, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 185
; LENGTH: 64415
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (9960)..(21146)
; OTHER INFORMATION: N = A, C G, or T/U
US-11-117-187-185

Query Match 19.5%; Score 30.6; DB 12; Length 64415;
Best Local Similarity 58.1%; Pred. No. 20;
Matches 54; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
Qy         65 CTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||| | ||     | || |||  | |||| ||||| |  | || |  | | || |
Db       7117 CTTGCTTCTTTAAGTATTATAGTATATTTGCTCCTAAACACTAAACCTAAACCCTACACC 7176

Qy        125 TTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              ||  || | ||  |  | || |  ||||| |||
Db       7177 TTAAATCCCAAACCCTAAATCTAATTCCTTAAC 7209
```

RESULT 4
US-11-121-086-13/c
; Sequence 13, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
; LENGTH: 191797
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-13

Query Match 19.4%; Score 30.4; DB 12; Length 191797;
Best Local Similarity 57.3%; Pred. No. 32;
Matches 55; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy         55 TCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTA 114
              || ||||   |   |||||| |||| |     |  | ||| | ||||||  |     ||
Db     162845 TCCTCCAGATCGCTCCTATTTTGAAATATATAAACGCTTACACACAAACACACTCATATA 162786

Qy        115 CTCATGCAACTTCCATGCAAACACGCACATATGTTT 150
                |||   ||| |||    | ||| ||||| | |||
Db     162785 TCCATTTTACTCCCACCACACCACACACATTTTTTT 162750
```

RESULT 5
US-09-925-065A-153052
; Sequence 153052, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 153052
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Homo sapiens

```
US-09-925-065A-153052

  Query Match             19.2%;  Score 30.2;  DB 6;  Length 436;
  Best Local Similarity   60.2%;  Pred. No. 6.1;
  Matches   50;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          4 GTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAG 63
              | | |||  | ||| | |   |||||||| |  | ||  | |||   ||| || |  ||
Db        256 GGCCTCTTGCTTGGAAGAGCCCATGCTTGGTTTAACGCCCTTCTGTCTCCGTCTTGAAAT 315

Qy         64 CCTTGCCTATTCTGAACCAAGAG 86
                | ||   ||| ||||| | |||
Db        316 TCATGATAATTTTGAACAAGGAG 338


RESULT 6
US-09-925-065A-285326
; Sequence 285326, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 285326
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-285326

  Query Match             19.0%;  Score 29.8;  DB 6;  Length 617;
  Best Local Similarity   60.5%;  Pred. No. 9;
  Matches   49;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy         74 TCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCA 133
              | || |  ||    |||| |||| ||||||||||||   |  |   |||||||  | |
Db        373 TGTGTAGAAATTTAATACATACTTCCAAACAATCCAGAGTTTTGTAGCAACTTATACTCC 432

Qy        134 AACACGCACATATGTTTCCTG 154
                ||  ||||  | |||   ||
Db        433 TACCTGCACTRAAGTTAAATG 453


RESULT 7
US-09-925-065A-285327
; Sequence 285327, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 285327
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-285327

  Query Match             19.0%;  Score 29.8;  DB 6;  Length 617;
  Best Local Similarity   60.5%;  Pred. No. 9;
  Matches   49;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy         74 TCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCA 133
              | || |  ||    |||| |||| ||||||||||||   |  |   |||||||  | |
Db        373 TGTGTAGAAATTTAATACATACTTCCAAACAATCCAGAGTTTTGTAGCAACTTATACTCC 432

Qy        134 AACACGCACATATGTTTCCTG 154
                ||  ||||  | |||   ||
Db        433 TACCTGCACTGAAGTTAAATG 453


RESULT 8
US-09-925-065A-285328
; Sequence 285328, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 285328
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-285328

  Query Match             19.0%;  Score 29.8;  DB 6;  Length 617;
  Best Local Similarity   60.5%;  Pred. No. 9;
  Matches   49;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy         74 TCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCA 133
              | || |  ||    |||| |||| ||||||||||||   |  |   |||||||  | |
Db        373 TGTGTAGAAATTTAATACATACTTCCAAACAATCCAGAGTTTTGTAGCAACTTATACTCC 432

Qy        134 AACACGCACATATGTTTCCTG 154
                ||  ||||  | |||   ||
Db        433 TACCTGCACTGAAGTTAAATG 453


RESULT 9
US-09-925-065A-12385
; Sequence 12385, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12385
;   LENGTH: 1363
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-12385

  Query Match             19.0%;  Score 29.8;  DB 6;  Length 1363;
  Best Local Similarity   50.3%;  Pred. No. 11;
  Matches   73;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          6 CATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCC 65
              || | ||| ||||  | ||   | ||  | |||| ||     |||   || || |
Db        765 CACCGCACCTGGCCAAGTATCAGTTTTCTTAACCATTGTCACTGTCTTATGGTTCTTCTG 824

Qy         66 TTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACT 125
              |    | |||||   || ||| ||    |     |||  |||    |   | |||||  |
Db        825 TGTTGTCTTCTGTGGCATGAGTATGTACAGGAAAAAAAGATCGGCTTATATAATGCATTT 884

Qy        126 TCCATGCAAACACGCACATATGTTT 150
                |||  || |  ||  || || |
Db        885 GAGATGGGAAAATTCAAGTACGTGT 909


RESULT 10
US-11-117-187-209/c
; Sequence 209, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
;  APPLICANT: PREUSS, DAPHNE
;  APPLICANT:  COPENHAVER, GREGORY
;  TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
;  FILE REFERENCE: ARCD:309US
;  CURRENT APPLICATION NUMBER: US/11/117,187
;  CURRENT FILING DATE:  2005-04-28
;  PRIOR APPLICATION NUMBER: US/09/531,120
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/125,219
;  PRIOR FILING DATE: 1999-03-18
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 209
;   LENGTH: 611587
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-11-117-187-209

  Query Match             18.7%;  Score 29.4;  DB 12;  Length 611587;
  Best Local Similarity   52.0%;  Pred. No. 91;
  Matches   66;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy          8 TCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTT 67
              |||| | || ||  | |   | ||  |||  | || ||      |   || | |  ||
Db     270144 TCTCTCCTGCCAAGCGAATGGATTTGTCATGCATTTGTTGCTAACTGCCGACACAAGCTG 270085

Qy         68 GCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTC 127
                || | ||  |||  ||     |  || || |  |  |||    |||||||   | ||| |
Db     270084 GCTTGTTAGGAAAAAAACCCTTTTCTTCTTCTTATGAATGTTGCTTACTCTGTCCACTGC 270025

Qy        128 CATGCAA 134
                || |||
Db     270024 GATACAA 270018


RESULT 11
US-09-925-065A-805392
; Sequence 805392, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 805392
;   LENGTH: 544
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-805392

  Query Match             18.3%;  Score 28.8;  DB 6;  Length 544;
  Best Local Similarity   69.6%;  Pred. No. 18;
  Matches   39;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         71 TATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTT 126
              ||||  |||  ||||  | ||| |   ||||| |||||| ||| |  |||||| ||
Db         23 TATTAAGAAAAAAGAAAAAACCAAAAACCAAAAAATCCAACTTTCCAATGCAAATT 78


RESULT 12
US-10-750-185-47517
; Sequence 47517, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47517
;   LENGTH: 1370
;   TYPE: DNA
;   ORGANISM: Bovine  19866880570548
US-10-750-185-47517

  Query Match             18.3%;  Score 28.8;  DB 8;  Length 1370;
```

```
  Best Local Similarity   54.8%;  Pred. No. 24;
  Matches   57;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          47 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 106
               ||||| | || | |  |||   || |||| |   | ||| | |    ||||  |||  ||
Db         660 TTGTTTCCTCATGCTTGCCGCTCCCATTCAGGGACTAGAAGCTGAAAACTCTGAAATTAT 719

Qy         107 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTT 150
               || | |    |  | || ||     || | | ||| |   ||||
Db         720 CCTTGTGCAACTAGGAAATTGTGGACAGAAATGCAAAATAGTTT 763


RESULT 13
US-10-750-623-47517
; Sequence 47517, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47517
;   LENGTH: 1370
;   TYPE: DNA
;   ORGANISM: Bovine  19866880570548
US-10-750-623-47517

  Query Match             18.3%;  Score 28.8;  DB 8;  Length 1370;
  Best Local Similarity   54.8%;  Pred. No. 24;
  Matches   57;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          47 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 106
               ||||| | || | |  |||   || |||| |   | ||| | |    ||||  |||  ||
Db         660 TTGTTTCCTCATGCTTGCCGCTCCCATTCAGGGACTAGAAGCTGAAAACTCTGAAATTAT 719

Qy         107 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTT 150
               || | |    |  | || ||     || | | ||| |   ||||
Db         720 CCTTGTGCAACTAGGAAATTGTGGACAGAAATGCAAAATAGTTT 763


RESULT 14
US-09-925-065A-845392
; Sequence 845392, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 845392
;   LENGTH: 628
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-845392

  Query Match             18.1%;  Score 28.4;  DB 6;  Length 628;
  Best Local Similarity   54.9%;  Pred. No. 25;
  Matches   56;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy          30 TTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGAT 89
               | ||| |   | | |   | ||    || |  ||| ||||  |||   ||    | || |
Db         495 TGGTGCTGGCCTCCAGATAGTGAGTTCTTGTAAGACTTGGTCATTTAAAAGTGTGCGGCT 554

Qy          90 ACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
               || ||| ||  ||  ||  |||| ||| |||   | |||||
Db         555 CCCCACTGCCCCACTCTCTCTCTTGCTCCTGCCTTTGCCATG 596


RESULT 15
US-09-925-065A-33976/c
; Sequence 33976, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33976
;   LENGTH: 1691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-33976

  Query Match             18.1%;  Score 28.4;  DB 6;  Length 1691;
  Best Local Similarity   54.9%;  Pred. No. 34;
  Matches   56;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy          30 TTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGAT 89
               | ||| |   | | |   | ||    || |  ||| ||||  |||   ||    | || |
Db        1191 TGGTGCTGGCCTCCAGATAGTGAGTTCTTGTAAGACTTGGTCATTTAAAAGTGTGCGGCT 1132

Qy          90 ACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
               || ||| ||  ||  ||  |||| ||| |||   | |||||
Db        1131 CCCCACTGCCCCACTCTCTCTCTTGCTCCTGCCTTTGCCATG 1090


Search completed: March  6, 2006, 10:32:51
Job time : 165.89 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:08:23 ; Search time 19.9642 Seconds
(without alignments)
10014.946 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208
Perfect score: 30
Sequence: 1 catgcttgttcaaccgttcgtcttgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30	100.0	30	5	AAH76335	Aah76335 Z. mays M
	2	30	100.0	255	5	AAH76340	Aah76340 Z. mays M
	3	30	100.0	1394	2	AAX07408	Aax07408 Zea mays
	4	30	100.0	1394	2	AAX07409	Aax07409 Zea mays
	5	30	100.0	1394	5	AAH76332	Aah76332 Z. mays M
	6	30	100.0	1394	5	AAH76333	Aah76333 Z. mays M
	7	24.2	80.7	158	5	AAH76334	Aah76334 Z. mays M
c	8	22	73.3	1304	6	ABQ75351	Abq75351 Human lun
	9	20.4	68.0	1542	6	ABZ13819	Abz13819 Arabidops
	10	20.2	67.3	2397	8	ACA44013	Aca44013 Prokaryot
	11	20	66.7	2000	11	ACL37526	Acl37526 Rice stre
	12	20	66.7	2000	11	ACL34976	Acl34976 Rice stre
	13	20	66.7	2000	12	ADJ40785	Adj40785 Plant cDN
	14	19.6	65.3	1735	11	ACL35519	Acl35519 Rice stre
	15	19.6	65.3	2000	8	ADA72061	Ada72061 Rice gene
c	16	19.6	65.3	2000	8	ADA71994	Ada71994 Rice gene
c	17	19.6	65.3	2000	8	ADA71736	Ada71736 Rice gene
c	18	19.6	65.3	2000	8	ADA72530	Ada72530 Rice gene
c	19	19.6	65.3	2000	11	ACL37405	Acl37405 Rice stre
c	20	19.6	65.3	2000	12	ADJ40977	Adj40977 Plant cDN
	21	19.4	64.7	2240	2	AAQ47668	Aaq47668 Sequence
	22	19.4	64.7	2243	4	AAF29956	Aaf29956 Rat brain
	23	19.4	64.7	2243	9	ACD26214	Acd26214 cDNA enco
	24	19.4	64.7	349980	6	ABQ81845	Abq81845 Bifidobac
c	25	19.2	64.0	2000	8	ADA73340	Ada73340 Rice gene
c	26	19	63.3	1324	8	ADA71777	Ada71777 Rice gene
	27	19	63.3	1388	3	AAC44356	Aac44356 Arabidops
c	28	19	63.3	1821	10	ADE93879	Ade93879 Ixodes sc
	29	19	63.3	2000	11	ACL35527	Acl35527 Rice stre
c	30	19	63.3	5671	4	AAK90190	Aak90190 Human dig
	31	19	63.3	5671	5	ABA20915	Aba20915 Human ner
c	32	19	63.3	5671	5	AAS39832	Aas39832 Genomic s
c	33	19	63.3	5671	9	ADB32792	Adb32792 Human nov
	34	19	63.3	198522	11	ACN44010	Acn44010 Human gen
c	35	18.8	62.7	883	6	ABK65288	Abk65288 Arabidops
c	36	18.8	62.7	883	10	ADD30938	Add30938 Plant yie
c	37	18.8	62.7	883	10	ADE37244	Ade37244 Plant yie
c	38	18.8	62.7	883	12	ADI41918	Adi41918 Plant tra
c	39	18.8	62.7	1083	4	AAS53644	Aas53644 Helicobac
c	40	18.8	62.7	1083	8	ACA34769	Aca34769 Prokaryot
c	41	18.8	62.7	1899	10	ABX06821	Abx06821 S. pneumo
c	42	18.8	62.7	1902	4	AAS55583	Aas55583 Streptoco
c	43	18.8	62.7	1902	8	ACA49918	Aca49918 Prokaryot
c	44	18.8	62.7	1908	12	ADJ63835	Adj63835 Plant lip
	45	18.8	62.7	2000	11	ACL35067	Acl35067 Rice stre

ALIGNMENTS

RESULT 1
AAH76335
ID AAH76335 standard; DNA; 30 BP.
XX
AC AAH76335;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The

```
CC   present sequence represents a DNA fragment -152 to -181 bases upstream of
CC   the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC   nucleotide sequence
XX
SQ   Sequence 30 BP; 4 A; 9 C; 5 G; 12 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 30;  DB 5;  Length 30;
  Best Local Similarity  100.0%;  Pred. No. 0.0045;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||||||
Db          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30


RESULT 2
AAH76340
ID   AAH76340 standard; DNA; 255 BP.
XX
AC   AAH76340;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 promoter fragment.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; promoter; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Example 5; Fig 8; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a Z. mays Ms45 promoter fragment
XX
SQ   Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 30;  DB 5;  Length 255;
  Best Local Similarity  100.0%;  Pred. No. 0.0058;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||||||
Db         39 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 68


RESULT 3
AAX07408
ID   AAX07408 standard; DNA; 1394 BP.
XX
AC   AAX07408;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW   Ms45; male; tissue-preferred; regulatory region; plant cells;
KW   plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS   Zea mays.
XX
PN   WO9859061-A1.
XX
PD   30-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US012895.
XX
PR   23-JUN-1997;   97US-00880499.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman GA,  Kendall TL;
XX
DR   WPI; 1999-105628/09.
XX
PT   New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT   - useful in mediating plant fertility, especially hybrid seed production.
XX
PS   Claim 2; Page 22-23; 39pp; English.
XX
CC   The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC   region. It may be used in the construction of a vector for a method of
CC   producing exogenous genes in a male tissue- preferred manner, which is
CC   useful in restoring or conferring fertility, such as in hybrid seed
CC   production. In conferring fertility, a monocot/dicot plant is transformed
CC   with the exogenous nucleotide sequence (a male sterility gene, preferably
CC   Ms45), which encodes a product selected from auxins, rolB and diptheria
CC   toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 30;  DB 2;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 0.007;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||||||
Db       1179 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 1208


RESULT 4
AAX07409
ID   AAX07409 standard; DNA; 1394 BP.
XX
AC   AAX07409;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW   Ms45; male; tissue-preferred; regulatory region; plant cells;
KW   plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS   Zea mays.
XX
PN   WO9859061-A1.
XX
PD   30-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US012895.
XX
```

PR 23-JUN-1997; 97US-00880499.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
DR WPI; 1999-105628/09.
XX
PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
PS Claim 3; Page 23-24; 39pp; English.
XX
CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue- preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
||||||||||||||||||||||||||||||
Db 1179 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 1208

RESULT 5
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
||||||||||||||||||||||||||||||
Db 1179 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 1208

RESULT 6
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
XX
AC AAH76333;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
||||||||||||||||||||||||||||||
Db 1179 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 1208

```
RESULT 7
AAH76334
ID   AAH76334 standard; DNA; 158 BP.
XX
AC   AAH76334;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 5; Page 47; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a DNA fragment -38 to -195 bases upstream of
CC   the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC   nucleotide sequence
XX
SQ   Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

  Query Match             80.7%;  Score 24.2;  DB 5;  Length 158;
  Best Local Similarity   89.7%;  Pred. No. 1.6;
  Matches   26;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCC 29
              |||||||||||||||||||||||||  | |
Db         25 CATGCTTGTTCAACCGTTCGTCTTTGTTC 53


RESULT 8
ABQ75351/c
ID   ABQ75351 standard; cDNA; 1304 BP.
XX
AC   ABQ75351;
XX
DT   05-NOV-2002  (first entry)
XX
DE   Human lung specific nucleic acid sequence SEQ ID NO:90.
XX
KW   Human; lung; lung specific nucleic acid; LSNA; lung specific protein;
KW   LSP; cytostatic; gene therapy; vaccine; metastasis; lung cancer;
KW   squamous cell carcinoma; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200264788-A2.
XX
PD   22-AUG-2002.
XX
PF   20-NOV-2001; 2001WO-US045080.
XX
PR   20-NOV-2000; 2000US-0252054P.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Macina RA,  Recipon H,  Chen S,  Sun Y,  Liu C;
XX
DR   WPI; 2002-657601/70.
XX
PT   New lung specific nucleic acid useful in gene therapy or as vaccines for
PT   treating lung cancer (e.g. squamous cell carcinoma) or non-cancerous lung
PT   diseases, as well as for diagnosing, monitoring or staging these
PT   diseases.
XX
PS   Claim 1; Page 206; 282pp; English.
XX
CC   The present invention describes an isolated lung specific nucleic acid
CC   (LSNA) comprising a sequence that: (a) encodes any of the 93 amino acid
CC   sequences comprising 17 - 733 amino acids, given in ABP52873 to ABP52965;
CC   (b) comprises any of 115 sequences comprising 148 - 3193 base pairs (bp),
CC   given in ABQ75262 to ABQ75376; (c) selectively hybridises to (a) or (b);
CC   or (d) has 60 % sequence identity to (a) or (b). LSNA and lung specific
CC   protein (LSP) sequences have cytostatic activity and can be used in gene
CC   therapy and vaccines. LSNAs and LSPs are useful for diagnosing and
CC   monitoring the presence and metastases of lung cancer in a patient. An
CC   antibody that specifically binds to an LSP can be used for determining
CC   the presence of an LSP in a sample, as well as for treating a patient
CC   with lung cancer, particularly by inducing an immune response against the
CC   lung cancer cell expressing the LSNAs or LSPs. In particular, these LSNAs
CC   and LSPs are useful for identifying, diagnosing, monitoring, staging,
CC   imaging and treating lung cancer (e.g. squamous cell carcinoma) and non-
CC   cancerous disease states in lung
XX
SQ   Sequence 1304 BP; 352 A; 347 C; 314 G; 291 T; 0 U; 0 Other;

  Query Match             73.3%;  Score 22;  DB 6;  Length 1304;
  Best Local Similarity   83.3%;  Pred. No. 18;
  Matches   25;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              | | |||||||  ||||||  ||||||||||
Db       1155 CTTTCTTGTTCCCCCGTTCCTCTTGTTCCA 1126


RESULT 9
ABZ13819
ID   ABZ13819 standard; DNA; 1542 BP.
XX
AC   ABZ13819;
XX
DT   21-JAN-2003  (first entry)
XX
DE   Arabidopsis thaliana stress regulated gene SEQ ID NO 1624.
XX
KW   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200216655-A2.
XX
PD   28-FEB-2002.
XX
PF   24-AUG-2001; 2001WO-US026685.
XX
PR   24-AUG-2000; 2000US-0227866P.
PR   26-JAN-2001; 2001US-0264647P.
PR   22-JUN-2001; 2001US-0300111P.
XX
```

PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 1624; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1542 BP; 345 A; 382 C; 370 G; 445 T; 0 U; 0 Other;

Query Match 68.0%; Score 20.4; DB 6; Length 1542;
Best Local Similarity 80.0%; Pred. No. 91;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
|||| |||||||||| ||||| |||||
Db 237 CATGATTGTTCAACCAATCGTCGGTTTCCA 266

RESULT 10
ACA44013
ID ACA44013 standard; DNA; 2397 BP.
XX
AC ACA44013;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #25670.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Pseudomonas putida.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU40143.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 31883; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2397 BP; 415 A; 825 C; 781 G; 376 T; 0 U; 0 Other;

Query Match 67.3%; Score 20.2; DB 8; Length 2397;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTG 25
|||||| || ||||||||||| ||||
Db 516 CATGCTGGTACAACCGTTCGGCTTG 540

RESULT 11
ACL37526
ID ACL37526 standard; cDNA; 2000 BP.
XX
AC ACL37526;
XX
DT 02-JUN-2005 (first entry)
XX
DE Rice stress-regulated promoter SEQ ID NO:16089.
XX
KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW agriculture.
XX
OS Oryza sativa.
XX
PN WO2003008540-A2.
XX
PD 30-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US019668.
XX
PR 22-JUN-2001; 2001US-0300112P.
PR 24-AUG-2001; 2001US-0314662P.
PR 26-SEP-2001; 2001US-0325277P.
PR 21-NOV-2001; 2001US-0332132P.
XX

```
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Kreps J,  Briggs SP,  Cooper B,  Glazebrook J,  Goff SA,  Katagiri F;
PI   Moughamer T,  Provart N,  Ricke D,  Zhu T;
XX
DR   WPI; 2003-248011/24.
XX
PT   New stress-responsive nucleic acid, useful for altering the
PT   responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT   stress, salt stress or osmotic stress.
XX
PS   Claim 48; SEQ ID NO 16089; 89pp; English.
XX
CC   The invention relates to novel abiotic stress responsive polynucleotides
CC   and polypeptides. Also disclosed are vectors, expression cassettes, host
CC   cells, and plants containing such polynucleotides. Also disclosed are
CC   methods for using the polynucleotides and polypeptides to alter the
CC   responsiveness of a plant to abiotic stress. The invention is useful in
CC   agriculture. The nucleic acid is useful for determining whether a test
CC   plant has been exposed to an abiotic stress condition. It is also useful
CC   for selecting an agent that alters abiotic stress regulated
CC   polynucleotide expression in a plant cell, and to identify a homolog or
CC   ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC   molecule and the polypeptide encoded by it are useful in altering the
CC   responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC   stress, osmotic stress or any of their combinations. The present sequence
CC   is used in the exemplification of the invention
XX
SQ   Sequence 2000 BP; 607 A; 443 C; 399 G; 549 T; 0 U; 2 Other;

  Query Match             66.7%;  Score 20;  DB 11;  Length 2000;
  Best Local Similarity   82.1%;  Pred. No. 1.4e+02;
  Matches   23;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 TGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||| ||||  |||||||||||| ||| |
Db       1576 TGCATGTTTGACCGTTCGTCTTATTCAA 1603


RESULT 12
ACL34976
ID   ACL34976 standard; cDNA; 2000 BP.
XX
AC   ACL34976;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Rice stress-regulated promoter SEQ ID NO:13539.
XX
KW   ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW   agriculture.
XX
OS   Oryza sativa.
XX
PN   WO2003008540-A2.
XX
PD   30-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019668.
XX
PR   22-JUN-2001; 2001US-0300112P.
PR   24-AUG-2001; 2001US-0314662P.
PR   26-SEP-2001; 2001US-0325277P.
PR   21-NOV-2001; 2001US-0332132P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Kreps J,  Briggs SP,  Cooper B,  Glazebrook J,  Goff SA,  Katagiri F;
PI   Moughamer T,  Provart N,  Ricke D,  Zhu T;
XX
DR   WPI; 2003-248011/24.
XX
PT   New stress-responsive nucleic acid, useful for altering the
PT   responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT   stress, salt stress or osmotic stress.
XX
PS   Claim 48; SEQ ID NO 13539; 89pp; English.
XX
CC   The invention relates to novel abiotic stress responsive polynucleotides
CC   and polypeptides. Also disclosed are vectors, expression cassettes, host
CC   cells, and plants containing such polynucleotides. Also disclosed are
CC   methods for using the polynucleotides and polypeptides to alter the
CC   responsiveness of a plant to abiotic stress. The invention is useful in
CC   agriculture. The nucleic acid is useful for determining whether a test
CC   plant has been exposed to an abiotic stress condition. It is also useful
CC   for selecting an agent that alters abiotic stress regulated
CC   polynucleotide expression in a plant cell, and to identify a homolog or
CC   ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC   molecule and the polypeptide encoded by it are useful in altering the
CC   responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC   stress, osmotic stress or any of their combinations. The present sequence
CC   is used in the exemplification of the invention
XX
SQ   Sequence 2000 BP; 552 A; 438 C; 444 G; 566 T; 0 U; 0 Other;

  Query Match             66.7%;  Score 20;  DB 11;  Length 2000;
  Best Local Similarity   82.1%;  Pred. No. 1.4e+02;
  Matches   23;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 TGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              | | |||| |||||||||||| ||| |
Db       1179 TACATGTTTAACCGTTCGTCTTATTCAA 1206


RESULT 13
ADJ40785
ID   ADJ40785 standard; cDNA; 2000 BP.
XX
AC   ADJ40785;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Plant cDNA #1785.
XX
KW   Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW   soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW   maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW   stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW   plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW   antifungal.
XX
OS   Eukaryota.
XX
PN   US2004016025-A1.
XX
PD   22-JAN-2004.
XX
PF   26-SEP-2002; 2002US-00260238.
XX
PR   26-SEP-2001; 2001US-0325277P.
PR   26-SEP-2001; 2001US-0325448P.
PR   04-APR-2002; 2002US-0370620P.
XX
PA   (BUDW/) BUDWORTH P.
PA   (MOUG/) MOUGHAMER T.
PA   (BRIG/) BRIGGS S P.
PA   (COOP/) COOPER B.
PA   (GLAZ/) GLAZEBROOK J.
PA   (GOFF/) GOFF S A.
PA   (KATA/) KATAGIRI F.
PA   (KREP/) KREPS J.
PA   (PROV/) PROVART N.
PA   (RICK/) RICKE D.
PA   (ZHUT/) ZHU T.
XX
PI   Budworth P,  Moughamer T,  Briggs SP,  Cooper B,  Glazebrook J;
```

PI Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
DR WPI; 2004-190374/18.
XX
PT New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT or high nutritional value.
XX
PS Claim 26; SEQ ID NO 1785; 230pp; English.
XX
CC The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2000 BP; 552 A; 438 C; 444 G; 566 T; 0 U; 0 Other;

Query Match 66.7%; Score 20; DB 12; Length 2000;
Best Local Similarity 82.1%; Pred. No. 1.4e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 TGCTTGTTCAACCGTTCGTCTTGTTCCA 30
| | |||| |||||||||||||| ||| |
Db 1179 TACATGTTTAACCGTTCGTCTTATTCAA 1206

RESULT 14
ACL35519
ID ACL35519 standard; cDNA; 1735 BP.
XX
AC ACL35519;
XX
DT 02-JUN-2005 (first entry)
XX
DE Rice stress-regulated promoter SEQ ID NO:14082.
XX
KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW agriculture.
XX
OS Oryza sativa.
XX
PN WO2003008540-A2.
XX
PD 30-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US019668.
XX
PR 22-JUN-2001; 2001US-0300112P.
PR 24-AUG-2001; 2001US-0314662P.
PR 26-SEP-2001; 2001US-0325277P.
PR 21-NOV-2001; 2001US-0332132P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
PI Moughamer T, Provart N, Ricke D, Zhu T;
XX
DR WPI; 2003-248011/24.
XX
PT New stress-responsive nucleic acid, useful for altering the
PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT stress, salt stress or osmotic stress.
XX
PS Claim 48; SEQ ID NO 14082; 89pp; English.
XX
CC The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention
XX
SQ Sequence 1735 BP; 502 A; 368 C; 279 G; 581 T; 0 U; 5 Other;

Query Match 65.3%; Score 19.6; DB 11; Length 1735;
Best Local Similarity 84.6%; Pred. No. 2e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 CTTGTTCAACCGTTCGTCTTGTTCCA 30
| |||| |||||||||||||| ||| |
Db 394 CATGTTTAACCGTTCGTCTTATTCAA 419

RESULT 15
ADA72061
ID ADA72061 standard; DNA; 2000 BP.
XX
AC ADA72061;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5386.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5386; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in

CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 605 A; 377 C; 326 G; 690 T; 0 U; 2 Other;

Query Match 65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity 84.6%; Pred. No. 2.1e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          5 CTTGTTCAACCGTTCGTCTTGTTCCA 30
              | |||| |||||||||||||| ||| |
Db        974 CATGTTTAACCGTTCGTCTTATTCAA 999
```

Search completed: March 5, 2006, 18:11:42
Job time : 19.9642 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 18:11:59 ; Search time 141.632 Seconds
(without alignments)
9910.279 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208
Perfect score: 30
Sequence: 1 catgcttgttcaaccgttcgtcttgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	30	100.0	687	9	CC656939	CC656939 OGWDQ20TV
	2	30	100.0	915	10	CG224225	CG224225 OG1AG08TV
	3	30	100.0	963	9	CC656933	CC656933 OGWDQ20TM
	4	27.4	91.3	296	10	CW445575	CW445575 fsbb001f1
c	5	22.2	74.0	308	10	CG700474	CG700474 ZMMBBc017
c	6	21.2	70.7	534	5	BU035739	BU035739 QHJ7N19.y
c	7	21.2	70.7	626	5	BU030489	BU030489 QHJ15I20.
	8	21.2	70.7	792	7	CN137013	CN137013 OX1_54_E0
c	9	21	70.0	154	1	AA811551	AA811551 oa99e03.s
c	10	21	70.0	351	11	CR192459	CR192459 Reverse s
c	11	21	70.0	571	11	CR064270	CR064270 Reverse s
c	12	21	70.0	599	11	CR233964	CR233964 Reverse s
c	13	21	70.0	769	11	CR188179	CR188179 Reverse s
c	14	21	70.0	832	9	CC848693	CC848693 NDL.2I14.
c	15	21	70.0	835	9	CC860012	CC860012 NDL.2I13.
	16	20.8	69.3	586	7	CO853973	CO853973 LM_SH5_00
	17	20.6	68.7	412	8	R12648	R12648 yf57f08.r1
	18	20.6	68.7	450	8	R12618	R12618 yf57b08.r1
c	19	20.6	68.7	582	1	AU290776	AU290776 AU290776
c	20	20.6	68.7	594	8	DR437963	DR437963 ACAC-aab6
	21	20.6	68.7	681	9	AZ817790	AZ817790 2M0087O22
c	22	20.6	68.7	1144	9	CC271930	CC271930 CH261-112
	23	20.4	68.0	353	9	AQ261553	AQ261553 CITBI-E1-
	24	20.4	68.0	373	3	BP857382	BP857382 BP857382
	25	20.4	68.0	396	2	BE522930	BE522930 M30D4STM
	26	20.4	68.0	398	3	BP836596	BP836596 BP836596
	27	20.4	68.0	514	2	BG354154	BG354154 00661 lea
	28	20.4	68.0	515	3	BP561510	BP561510 BP561510
	29	20.4	68.0	536	1	AV439701	AV439701 AV439701
	30	20.4	68.0	559	1	AV825937	AV825937 AV825937
c	31	20.4	68.0	560	6	CA230882	CA230882 SCJFFL3C0
c	32	20.4	68.0	590	7	CV153842	CV153842 Eucinetus
	33	20.4	68.0	598	10	CW788317	CW788317 SP_Ba004
c	34	20.4	68.0	600	9	BH876120	BH876120 hr29b04.g
	35	20.4	68.0	602	10	CL583422	CL583422 OB_Ba007
	36	20.4	68.0	603	9	BZ694195	BZ694195 SP_Ba004
	37	20.4	68.0	615	9	AQ453944	AQ453944 HS_5165_B
c	38	20.4	68.0	639	7	CV153844	CV153844 Eucinetus
	39	20.4	68.0	668	9	AQ260121	AQ260121 CITBI-E1-
	40	20.4	68.0	779	6	CD825116	CD825116 BN25.059M
	41	20.4	68.0	807	6	CD836362	CD836362 BN45.048I
c	42	20.4	68.0	836	9	BH480110	BH480110 BOGHY44TR
c	43	20.4	68.0	967	9	BZ684538	BZ684538 PUBGO31TD
c	44	20.4	68.0	1205	9	B10570	B10570 T15J6-T7 TA
c	45	20.2	67.3	431	8	R32995	R32995 yh73g09.r1

ALIGNMENTS

RESULT 1
CC656939/c
LOCUS CC656939 687 bp DNA linear GSS 19-JUN-2003
DEFINITION OGWDQ20TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15, genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 687)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGWDQ20TM
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: methylation filtered.
FEATURES Location/Qualifiers
source 1. .687
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0554D15"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 100.0%; Score 30; DB 9; Length 687;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30

```
                    ||||||||||||||||||||||||||||||
Db        397 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 368


RESULT 2
CG224225
LOCUS       CG224225                915 bp    DNA     linear   GSS 22-AUG-2003
DEFINITION  OG1AG08TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0716B15,
            genomic survey sequence.
ACCESSION   CG224225
VERSION     CG224225.1  GI:34124113
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1AG08TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0716B15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 10;  Length 915;
  Best Local Similarity   100.0%;  Pred. No. 0.026;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||||||
Db        558 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 587


RESULT 3
CC656933
LOCUS       CC656933                963 bp    DNA     linear   GSS 19-JUN-2003
DEFINITION  OGWDQ20TM ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15,
            genomic survey sequence.
ACCESSION   CC656933
VERSION     CC656933.1  GI:32060225
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 963)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OGWDQ20TV
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .963
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0554D15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             100.0%;  Score 30;  DB 9;  Length 963;
  Best Local Similarity   100.0%;  Pred. No. 0.026;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||||||
Db        776 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 805


RESULT 4
CW445575
LOCUS       CW445575                296 bp    DNA     linear   GSS 02-NOV-2004
DEFINITION  fsbb001f170m16k0 Sorghum methylation filtered library (LibID: 104)
            Sorghum bicolor genomic clone fsbb001f170m16, genomic survey
            sequence.
ACCESSION   CW445575
VERSION     CW445575.1  GI:55193536
KEYWORDS    GSS.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 296)
  AUTHORS   Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
            Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
            McMenamy,J., Smith,M., Holeman,H., Roe,B.A, Wiley,G., Korf,I.F.,
            Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
            Martienssen,R.A.
  TITLE     Sorghum genome sequencing by methylation filtration
  JOURNAL   PLoS Biol. 3 (1), e13 (2005)
   PUBMED   15660154
COMMENT     Contact: Bedell JA
            Orion Genomics, LLC
            4041 Forest Park Ave, St. Louis, MO 63108, USA
            Tel: 314 615 6979
            Fax: 314 615 5975
            Email: jbedell@oriongenomics.com
            Plate: fsbb001f170  row: m  column: 16
            Seq primer: k Reverse
            Class: methylation filtered
            High quality sequence stop: 296.
FEATURES             Location/Qualifiers
     source          1. .296
                     /organism="Sorghum bicolor"
                     /mol_type="genomic DNA"
                     /cultivar="ATx623"
                     /db_xref="taxon:4558"
                     /clone="fsbb001f170m16"
                     /clone_lib="Sorghum methylation filtered library (LibID:
                     104)"
                     /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
                     prepared from purified nuclei was randomly sheared,
```

end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."

ORIGIN

Query Match 91.3%; Score 27.4; DB 10; Length 296;
Best Local Similarity 96.6%; Pred. No. 0.3;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||| |||
Db        140 ATGCTTGTTCAACCGTTCGTCTTGTACCA 168
```

RESULT 5
CG700474/c
LOCUS CG700474 308 bp DNA linear GSS 15-OCT-2003
DEFINITION ZMMBBc0174H15f ZMMBBc (EcoRI) Zea mays genomic clone ZMMBBc0174H15 5', genomic survey sequence.
ACCESSION CG700474
VERSION CG700474.1 GI:37688275
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 308)
AUTHORS Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C., Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE Sequencing of the maize genome at PGIR (2003c)
JOURNAL Unpublished (2003)
COMMENT Contact: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 57.
FEATURES Location/Qualifiers
source 1. .308
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBc0174H15"
/lab_host="E. coli DH10B"
/clone_lib="ZMMBBc (EcoRI)"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI"
ORIGIN

Query Match 74.0%; Score 22.2; DB 10; Length 308;
Best Local Similarity 88.9%; Pred. No. 60;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTT 27
              |||||| ||||| ||||||||||| ||
Db        226 CATGCTGGTTCAGCCGTTCGTCTTCTT 200
```

RESULT 6
BU035739/c
LOCUS BU035739 534 bp mRNA linear EST 23-AUG-2002
DEFINITION QHJ7N19.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA clone QHJ7N19, mRNA sequence.
ACCESSION BU035739
VERSION BU035739.1 GI:22471259
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 534)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project http://compgenomics.ucdavis.edu/
JOURNAL Unpublished (2002)
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QH_CA_Contig3046, see http://cgpdb.ucdavis.edu/ for details.
Plate: QHJ7 row: N column: 19.
FEATURES Location/Qualifiers
source 1. .534
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHJ7N19"
/lab_host="E.coli"
/clone_lib="QH_EFGHJ sunflower RHA280"
/note="Vector: pBRcDNASfiAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_TISSUE=flowers environmental stress
TAG_LIB=QH_EFGHJ sunflower RHA280
TAG_SEQ=CGAATGCGGG"
ORIGIN

Query Match 70.7%; Score 21.2; DB 5; Length 534;
Best Local Similarity 88.5%; Pred. No. 1.8e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTT 27
              |||||| ||||| |||||||||| ||
Db        104 ATGCTTCTTCAAGCGTTCGTCTTTTT 79
```

RESULT 7
BU030489/c
LOCUS BU030489 626 bp mRNA linear EST 23-AUG-2002
DEFINITION QHJ15I20.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA clone QHJ15I20, mRNA sequence.
ACCESSION BU030489
VERSION BU030489.1 GI:22466009
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 626)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J.,

```
            Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
            Lai,Z., Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QH_CA_Contig3046, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QHJ15  row: I  column: 20.
FEATURES             Location/Qualifiers
     source          1. .626
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA280"
                     /db_xref="taxon:4232"
                     /clone="QHJ15I20"
                     /lab_host="E.coli"
                     /clone_lib="QH_EFGHJ sunflower RHA280"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_TISSUE=flowers environmental stress
                     TAG_LIB=QH_EFGHJ sunflower RHA280
                     TAG_SEQ=CGAATGCGGG"
ORIGIN

  Query Match             70.7%;  Score 21.2;  DB 5;  Length 626;
  Best Local Similarity   88.5%;  Pred. No. 1.9e+02;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTT 27
              ||||||  |||||  ||||||||||  ||
Db        104 ATGCTTCTTCAAGCGTTCGTCTTTTT 79


RESULT 8
CN137013
LOCUS       CN137013                 792 bp    mRNA    linear   EST 01-APR-2004
DEFINITION  OX1_54_E01.g1_A002 Oxidatively-stressed leaves and roots Sorghum
            bicolor cDNA clone OX1_54_E01_A002 5', mRNA sequence.
ACCESSION   CN137013
VERSION     CN137013.1  GI:45970259
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C.,
            Sun,F., Sullivan,R., Lim,S., Eastman,A. and Pratt,L.H.
  TITLE     An EST database from Sorghum: oxidatively stressed leaves and roots
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: OX1_54_E01.b1_A002
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
            the Human Genome Center, University of Tokyo Institute of Medical
            Science; plant material and RNA prepared at Texas A & M University;
            sequencing done in the Laboratory for Genomics and Bioinformatics,
            University of Georgia. Sequence ends have been trimmed to exclude
            vector and regions below Phred quality 16. Three-prime sequences
            are presented as their reverse complement and have been trimmed to
            exclude polyA.
            Seq primer: Sug5 (CTTCTGCTCTAAAAGCTGCG).
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /cultivar="BTx623"
                     /db_xref="taxon:4558"
                     /clone="OX1_54_E01_A002"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Oxidatively-stressed leaves and roots"
                     /note="Organ: Leaf and Root; Vector: pME18S-FL3; Site_1:
                     XhoI; Site_2: XhoI; The library was prepared from polyA+
                     RNA from oxidatively stressed, hydroponically grown
                     sorghum seedlings. At 8 days of age, growth medium was
                     supplemented with hydrogen peroxide to 0.003% and leaves
                     were misted with 10 uM methyl viologen. Leaves and roots
                     were harvested at 3, 12 and 27 hr after treatment and all
                     tissue pooled. Double-stranded cDNA was cloned
                     unidirectionally into different DraIII sites of the
                     pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG,
                     3-prime DraIII site is CACCATGTG). XhoI excises the cDNA
                     insert."
ORIGIN

  Query Match             70.7%;  Score 21.2;  DB 7;  Length 792;
  Best Local Similarity   88.5%;  Pred. No. 1.9e+02;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 CTTGTTCAACCGTTCGTCTTGTTCCA 30
              |||  | ||||||||||||||| |||||
Db         17 CTTCTCCAACCGTTCGTCTTCTTCCA 42


RESULT 9
AA811551/c
LOCUS       AA811551                 154 bp    mRNA    linear   EST 19-FEB-1998
DEFINITION  oa99e03.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1320412 3'
            similar to SW:COPE_BOVIN Q28104 COATOMER EPSILON SUBUNIT ;, mRNA
            sequence.
ACCESSION   AA811551
VERSION     AA811551.1  GI:2881162
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 154)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
```

```
           Insert Length: 1232   Std Error: 0.00
           Seq primer: -40m13 fwd. ET from Amersham
           High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .154
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1320412"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GCB1"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     (5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ). Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN

  Query Match           70.0%;  Score 21;  DB 1;  Length 154;
  Best Local Similarity  82.8%;  Pred. No. 1.8e+02;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              | ||||||| |||| ||||| |||||||
Db         43 AAGCTTGTTAAACCCATCGTCCTGTTCCA 15


RESULT 10
CR192459/c
LOCUS       CR192459                351 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Reverse strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN146a06, genomic survey sequence.
ACCESSION   CR192459
VERSION     CR192459.1  GI:49971308
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 351)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .351
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN146a06"
                     /clone_lib="MHPN"
ORIGIN

  Query Match           70.0%;  Score 21;  DB 11;  Length 351;
  Best Local Similarity  82.8%;  Pred. No. 2.1e+02;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||| || || ||||||||  |||||||
Db        307 ATGCTGGTCCAGCCGTTCGTGATGTTCCA 279


RESULT 11
CR064270/c
LOCUS       CR064270                571 bp    DNA     linear   GSS 05-JUL-2004
DEFINITION  Reverse strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN190119, genomic survey sequence.
ACCESSION   CR064270
VERSION     CR064270.1  GI:49797742
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 571)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .571
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN190119"
                     /clone_lib="MHPN"
ORIGIN

  Query Match           70.0%;  Score 21;  DB 11;  Length 571;
  Best Local Similarity  82.8%;  Pred. No. 2.2e+02;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||| || || ||||||||  |||||||
Db        309 ATGCTGGTCCAGCCGTTCGTGATGTTCCA 281


RESULT 12
CR233964/c
LOCUS       CR233964                599 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Reverse strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN144h08, genomic survey sequence.
ACCESSION   CR233964
VERSION     CR233964.1  GI:50012813
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .599
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN144h08"
                     /clone_lib="MHPN"
ORIGIN

  Query Match           70.0%;  Score 21;  DB 11;  Length 599;
  Best Local Similarity  82.8%;  Pred. No. 2.3e+02;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
```

```
              ||||| || || ||||||||  |||||||
Db        310 ATGCTGGTCCAGCCGTTCGTGATGTTCCA 282


RESULT 13
CR188179/c
LOCUS       CR188179                769 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Reverse strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN234h03, genomic survey sequence.
ACCESSION   CR188179
VERSION     CR188179.1  GI:49967028
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .769
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN234h03"
                     /clone_lib="MHPN"
ORIGIN

  Query Match             70.0%;  Score 21;  DB 11;  Length 769;
  Best Local Similarity   82.8%;  Pred. No. 2.3e+02;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||| || || ||||||||  |||||||
Db        307 ATGCTGGTCCAGCCGTTCGTGATGTTCCA 279


RESULT 14
CC848693/c
LOCUS       CC848693                832 bp    DNA     linear   GSS 24-JUL-2003
DEFINITION  NDL.2I14.T7 Notre Dame Liverpool Aedes aegypti genomic clone
            NotreDame Liverpool-2I14, genomic survey sequence.
ACCESSION   CC848693
VERSION     CC848693.1  GI:33197857
KEYWORDS    GSS.
SOURCE      Aedes aegypti (yellow fever mosquito)
  ORGANISM  Aedes aegypti
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Culicidae; Culicinae; Culicini; Aedes; Stegomyia.
REFERENCE   1  (bases 1 to 832)
  AUTHORS   Loftus,B., Shetty,J., Knudson,D. and Severson,D.
  TITLE     BAC end sequencing of Aedes aegypti
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: NDL.2I14.SP6
            Contact: Brendan Loftus
            Department of Eukaryotic Genomics
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-3543
            Fax: 301-838-0208
            Email: enta@tigr.org
            Library was provided by David Severson
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .832
                     /organism="Aedes aegypti"
                     /mol_type="genomic DNA"
                     /strain="Liverpool"
                     /db_xref="taxon:7159"
                     /clone="NotreDame Liverpool-2I14"
                     /clone_lib="Notre Dame Liverpool"
                     /note="Vector: pECBAC1; Site_1: Hind III; The library was
                     prepared from whole body tissue of newly hatched L1 larvae
                     by David Severson at the University of Notre Dame and
                     Hongbin Zhang"
ORIGIN

  Query Match             70.0%;  Score 21;  DB 9;  Length 832;
  Best Local Similarity   82.8%;  Pred. No. 2.4e+02;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              || |||||||||| || |||||| | ||
Db        242 ATCCTTGTTCAACCATTTGTCTTGATTCA 214


RESULT 15
CC860012/c
LOCUS       CC860012                835 bp    DNA     linear   GSS 24-JUL-2003
DEFINITION  NDL.2I13.T7 Notre Dame Liverpool Aedes aegypti genomic clone
            NotreDame Liverpool-2I13, genomic survey sequence.
ACCESSION   CC860012
VERSION     CC860012.1  GI:33220022
KEYWORDS    GSS.
SOURCE      Aedes aegypti (yellow fever mosquito)
  ORGANISM  Aedes aegypti
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Culicidae; Culicinae; Culicini; Aedes; Stegomyia.
REFERENCE   1  (bases 1 to 835)
  AUTHORS   Loftus,B., Shetty,J., Knudson,D. and Severson,D.
  TITLE     BAC end sequencing of Aedes aegypti
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: NDL.2I13.SP6
            Contact: Brendan Loftus
            Department of Eukaryotic Genomics
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-3543
            Fax: 301-838-0208
            Email: enta@tigr.org
            Library was provided by David Severson
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .835
                     /organism="Aedes aegypti"
                     /mol_type="genomic DNA"
                     /strain="Liverpool"
                     /db_xref="taxon:7159"
                     /clone="NotreDame Liverpool-2I13"
                     /clone_lib="Notre Dame Liverpool"
                     /note="Vector: pECBAC1; Site_1: Hind III; The library was
                     prepared from whole body tissue of newly hatched L1 larvae
                     by David Severson at the University of Notre Dame and
                     Hongbin Zhang"
ORIGIN

  Query Match             70.0%;  Score 21;  DB 9;  Length 835;
  Best Local Similarity   82.8%;  Pred. No. 2.4e+02;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              || |||||||||| || |||||| | ||
Db        242 ATCCTTGTTCAACCATTTGTCTTGATTCA 214


Search completed: March  6, 2006, 01:57:49
Job time : 141.632 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 21:54:12 ; Search time 8.38752 Seconds
(without alignments)
6357.883 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208
Perfect score: 30
Sequence: 1 catgcttgttcaaccgttcgtcttgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		30	100.0	1394	3	US-08-880-499-1	Sequence 1, Appli
2		30	100.0	1394	3	US-08-880-499-2	Sequence 2, Appli
3		19.4	64.7	2243	2	US-07-937-609-15	Sequence 15, Appl
4		19.4	64.7	2243	3	US-08-029-170-15	Sequence 15, Appl
5		19.4	64.7	2243	3	US-09-443-745-15	Sequence 15, Appl
6		19	63.3	45571	3	US-09-949-016-16262	Sequence 16262, A
c 7		18.8	62.7	7812	3	US-09-368-590-1	Sequence 1, Appli
c 8		18.8	62.7	8756	3	US-09-949-016-1438	Sequence 1438, Ap
9		18.8	62.7	9541	3	US-08-961-527-132	Sequence 132, App
c 10		18.8	62.7	112874	3	US-09-949-016-13180	Sequence 13180, A
c 11		18.6	62.0	549	3	US-09-270-767-29480	Sequence 29480, A
c 12		18.6	62.0	1120	3	US-09-270-767-13492	Sequence 13492, A
13		18.4	61.3	1260	3	US-09-248-796A-3293	Sequence 3293, Ap
c 14		18.2	60.7	2710	3	US-09-573-080A-7	Sequence 7, Appli
c 15		18	60.0	601	3	US-09-949-016-200252	Sequence 200252,
c 16		18	60.0	601	3	US-09-949-016-200253	Sequence 200253,
17		18	60.0	889	3	US-08-956-171E-1190	Sequence 1190, Ap
18		18	60.0	889	3	US-08-781-986A-1190	Sequence 1190, Ap
c 19		18	60.0	1591	2	US-08-728-956-3	Sequence 3, Appli
c 20		18	60.0	33675	3	US-09-921-992-2	Sequence 2, Appli
21		18	60.0	111509	3	US-09-949-016-17379	Sequence 17379, A
c 22		18	60.0	254366	3	US-09-822-871-3	Sequence 3, Appli
23		17.8	59.3	393	3	US-09-583-110-1388	Sequence 1388, Ap
24		17.8	59.3	393	3	US-09-107-433-743	Sequence 743, App
c 25		17.8	59.3	499	3	US-09-270-767-1993	Sequence 1993, Ap
c 26		17.8	59.3	499	3	US-09-270-767-17275	Sequence 17275, A
27		17.8	59.3	601	3	US-09-949-016-142937	Sequence 142937,
c 28		17.8	59.3	601	3	US-09-949-016-204029	Sequence 204029,
c 29		17.8	59.3	601	3	US-09-949-016-204030	Sequence 204030,
c 30		17.8	59.3	978	3	US-08-858-207A-174	Sequence 174, App
31		17.8	59.3	999	3	US-08-961-527-193	Sequence 193, App
32		17.8	59.3	2009	2	US-07-958-222A-1	Sequence 1, Appli
c 33		17.8	59.3	8073	3	US-09-949-016-12327	Sequence 12327, A
c 34		17.8	59.3	8074	3	US-09-949-016-16645	Sequence 16645, A
35		17.8	59.3	9472	2	US-08-325-547-9	Sequence 9, Appli
36		17.8	59.3	40505	3	US-09-949-016-13439	Sequence 13439, A
37		17.8	59.3	72843	3	US-09-949-016-12574	Sequence 12574, A
38		17.8	59.3	141115	3	US-09-949-016-17490	Sequence 17490, A
39		17.8	59.3	183202	3	US-09-949-016-13614	Sequence 13614, A
c 40		17.8	59.3	265038	3	US-09-949-016-15779	Sequence 15779, A
41		17.8	59.3	276237	3	US-09-949-016-17504	Sequence 17504, A
c 42		17.4	58.0	331	3	US-09-640-211A-169	Sequence 169, App
43		17.4	58.0	601	3	US-09-949-016-178197	Sequence 178197,
44		17.4	58.0	719	3	US-09-107-433-820	Sequence 820, App
45		17.4	58.0	789	3	US-09-252-991A-11034	Sequence 11034, A

ALIGNMENTS

```
RESULT 1
US-08-880-499-1
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
;   APPLICANT: Albertson, Marc C.
;   APPLICANT: Fox, Tim W.
;   APPLICANT: Carl, Garnaat W.
;   APPLICANT: Huffman, Gary A.
;   APPLICANT: Kendall, Timmy L.
;   TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
;   TITLE OF INVENTION: AND METHOD OF USING SAME
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
;     STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
;     STREET: Box 1000
;     CITY: Johnston
;     STATE: Iowa
;     COUNTRY: USA
;     ZIP: 50131
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/880,499
;     FILING DATE: CONCURRENTLY HEREWITH
;     CLASSIFICATION: 800
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Sweeney, Patricia A.
;     REGISTRATION NUMBER: 32,733
;     REFERENCE/DOCKET NUMBER: 0578
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (515) 248-4800
;     TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1394 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-08-880-499-1

  Query Match 100.0%; Score 30; DB 3; Length 1394;
```

```
  Best Local Similarity   100.0%;  Pred. No. 0.0012;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||||||
Db       1179 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 1208


RESULT 2
US-08-880-499-2
; Sequence 2, Application US/08880499
; Patent No. 6037523
;  GENERAL INFORMATION:
;    APPLICANT:  Albertson, Marc C.
;    APPLICANT:  Fox, Tim W.
;    APPLICANT:  Carl, Garnaat W.
;    APPLICANT:  Huffman, Gary A.
;    APPLICANT:  Kendall, Timmy L.
;    TITLE OF INVENTION:  MALE TISSUE-PREFERRED REGULATORY REGION
;    TITLE OF INVENTION:  AND METHOD OF USING SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PIONEER HI-BRED INTERNATIONAL, INC.
;      STREET:  Darwin Building, 7100 N.W. 62nd Ave., P.O.
;      STREET:  Box 1000
;      CITY:  Johnston
;      STATE:  Iowa
;      COUNTRY:  USA
;      ZIP:  50131
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/880,499
;      FILING DATE:  CONCURRENTLY HEREWITH
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sweeney, Patricia A.
;      REGISTRATION NUMBER:  32,733
;      REFERENCE/DOCKET NUMBER:  0578
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 248-4800
;      TELEFAX:  (515) 248-4844
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1394 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-880-499-2

  Query Match             100.0%;  Score 30;  DB 3;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 0.0012;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||||||
Db       1179 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 1208


RESULT 3
US-07-937-609-15
; Sequence 15, Application US/07937609
; Patent No. 5319073
;  GENERAL INFORMATION:
;    APPLICANT:  WANK, Stephen A.
;    TITLE OF INVENTION:  CLONING AND FUNCTIONAL EXPRESSION OF
;    TITLE OF INVENTION:  CHOLECYSTOKININ RECEPTOR-ENCODING DNA
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/937,609
;      FILING DATE:  19920902
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/831,248
;      FILING DATE:  07-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/861,769
;      FILING DATE:  01-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/928,033
;      FILING DATE:  11-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  40399/166 NIHD
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2243 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  rat brain CCKB receptor
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  136..1494
US-07-937-609-15

  Query Match             64.7%;  Score 19.4;  DB 2;  Length 2243;
  Best Local Similarity   79.3%;  Pred. No. 55;
  Matches   23;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCC 29
              | ||||||  |||||  | ||||||||| |
Db          5 CCTGCTTGCTCAACTCTACGTCTTGTTTC 33


RESULT 4
US-08-029-170-15
; Sequence 15, Application US/08029170
; Patent No. 6169173
;  GENERAL INFORMATION:
;    APPLICANT:  WANK, Stephen A.
;    TITLE OF INVENTION:  CLONING AND FUNCTIONAL EXPRESSION OF
;    TITLE OF INVENTION:  CHOLECYSTOKININ RECEPTOR-ENCODING DNA
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/029,170
;       FILING DATE:  19930310
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/937,609
;       FILING DATE:  02-SEP-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/928,033
;       FILING DATE:  11-AUG-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/861,769
;       FILING DATE:  01-APR-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/831,248
;       FILING DATE:  07-FEB-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BENT, Stephen A.
;       REGISTRATION NUMBER:  29,768
;       REFERENCE/DOCKET NUMBER:  40399/166 NIHD
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703)836-9300
;       TELEFAX:  (703)683-4109
;       TELEX:  899149
;   INFORMATION FOR SEQ ID NO:  15:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2243 base pairs
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     IMMEDIATE SOURCE:
;       CLONE:  rat brain CCKB receptor
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  136..1494
US-08-029-170-15

  Query Match             64.7%;  Score 19.4;  DB 3;  Length 2243;
  Best Local Similarity   79.3%;  Pred. No. 55;
  Matches   23;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCC 29
              | |||||| |||||  | ||||||||| |
Db          5 CCTGCTTGCTCAACTCTACGTCTTGTTTC 33


RESULT 5
US-09-443-745-15
; Sequence 15, Application US/09443745
; Patent No. 6706493
;  GENERAL INFORMATION:
;    APPLICANT:  WANK, Stephen A.
;    TITLE OF INVENTION:  CLONING AND FUNCTIONAL EXPRESSION OF
;    TITLE OF INVENTION:  CHOLECYSTOKININ RECEPTOR-ENCODING DNA
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/443,745
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/029,170
;      FILING DATE:  10-MAR-1993
;      APPLICATION NUMBER:  US 07/937,609
;      FILING DATE:  02-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/928,033
;      FILING DATE:  11-AUG-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/861,769
;      FILING DATE:  01-APR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/831,248
;      FILING DATE:  07-FEB-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  40399/166 NIHD
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2243 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  rat brain CCKB receptor
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  136..1494
US-09-443-745-15

  Query Match             64.7%;  Score 19.4;  DB 3;  Length 2243;
  Best Local Similarity   79.3%;  Pred. No. 55;
  Matches   23;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCC 29
              | |||||| |||||  | ||||||||| |
Db          5 CCTGCTTGCTCAACTCTACGTCTTGTTTC 33


RESULT 6
US-09-949-016-16262
; Sequence 16262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT:  VENTER, J. Craig et al.
;  TITLE OF INVENTION:  POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE:  CL001307
;  CURRENT APPLICATION NUMBER:  US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER:  60/241,755
;  PRIOR FILING DATE:  2000-10-20
;  PRIOR APPLICATION NUMBER:  60/237,768
;  PRIOR FILING DATE:  2000-10-03
;  PRIOR APPLICATION NUMBER:  60/231,498
;  PRIOR FILING DATE:  2000-09-08
;  NUMBER OF SEQ ID NOS:  207012
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 16262
;   LENGTH:  45571
;   TYPE:  DNA
;   ORGANISM:  Human
US-09-949-016-16262

  Query Match             63.3%;  Score 19;  DB 3;  Length 45571;
```

Best Local Similarity 81.5%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          3 TGCTTGTTCAACCGTTCGTCTTGTTCC 29
              || ||||||| ||||| |||| ||||
Db      38119 TGTTTGTTCATTCGTTCTTCTTCTTCC 38145
```

RESULT 7
US-09-368-590-1/c
; Sequence 1, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7812
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6879)
; NAME/KEY: unsure
; LOCATION: (100)...(102)
; NAME/KEY: unsure
; LOCATION: (1021)...(1023)
; NAME/KEY: unsure
; LOCATION: (2266)...(2268)
US-09-368-590-1

Query Match 62.7%; Score 18.8; DB 3; Length 7812;
Best Local Similarity 76.7%; Pred. No. 1.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||| | |||| | ||||||   ||||
Db        841 CATGCTTCTGCAACAGGTCGTCTGCCTCCA 812
```

RESULT 8
US-09-949-016-1438/c
; Sequence 1438, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1438
; LENGTH: 8756
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1438

Query Match 62.7%; Score 18.8; DB 3; Length 8756;
Best Local Similarity 76.7%; Pred. No. 1.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||| | |||| | ||||||   ||||
Db       1912 CATGCTTCTGCAACAGGTCGTCTGCCTCCA 1883
```

RESULT 9
US-08-961-527-132
; Sequence 132, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9541 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-132

Query Match 62.7%; Score 18.8; DB 3; Length 9541;
Best Local Similarity 76.7%; Pred. No. 1.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              || |||||| ||| |||||   || |||||
Db       8049 CAGGCTTGTCCAAACGTTCCATTTTTTCCA 8078
```

RESULT 10
US-09-949-016-13180/c
; Sequence 13180, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13180
; LENGTH: 112874
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(112874)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13180

Query Match 62.7%; Score 18.8; DB 3; Length 112874;
Best Local Similarity 76.7%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
||||||| | |||| | |||||| ||||
Db 41167 CATGCTTCTGCAACAGGTCGTCTGCCTCCA 41138

RESULT 11
US-09-270-767-29480/c
; Sequence 29480, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29480
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-29480

Query Match 62.0%; Score 18.6; DB 3; Length 549;
Best Local Similarity 84.0%; Pred. No. 98;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 TGCTTGTTCAACCGTTCGTCTTGTT 27
||||||| || |||| |||||||
Db 223 TGCTTGTTGAAGAGTTCTTCTTGTT 199

RESULT 12
US-09-270-767-13492/c
; Sequence 13492, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13492
; LENGTH: 1120
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-13492

Query Match 62.0%; Score 18.6; DB 3; Length 1120;
Best Local Similarity 84.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 TGCTTGTTCAACCGTTCGTCTTGTT 27
||||||| || |||| |||||||
Db 223 TGCTTGTTGAAGAGTTCTTCTTGTT 199

RESULT 13
US-09-248-796A-3293
; Sequence 3293, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3293
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3293

Query Match 61.3%; Score 18.4; DB 3; Length 1260;
Best Local Similarity 78.6%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 TGCTTGTTCAACCGTTCGTCTTGTTCCA 30
||||||||||||| ||| || || ||| |||
Db 903 TGCTTGTTCAACTGTTGGTGTTCCGCCA 930

RESULT 14
US-09-573-080A-7/c
; Sequence 7, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATIN
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 2710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(2710)
; OTHER INFORMATION: charlie3
; NAME/KEY: misc_feature
; OTHER INFORMATION: n is a, c, g or t
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database for repetitive elements (repbase)
; DATABASE ENTRY DATE: ____-__-__
; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-7

Query Match 60.7%; Score 18.2; DB 3; Length 2710;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          3 TGCTTGTTCAACCGTTCGTCTTGTT 27
              ||||  |||| | |||||||||:|||
Db       1209 TGCTGTTTCACCAGTTCGTCTKGTT 1185
```

RESULT 15
US-09-949-016-200252/c
; Sequence 200252, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 200252
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-200252

Query Match 60.0%; Score 18; DB 3; Length 601;
Best Local Similarity 80.8%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          5 CTTGTTCAACCGTTCGTCTTGTTCCA 30
              |||||| ||| |||| | ||| ||||
Db        239 CTTGTTAAACAGTTCCTGTTGCTCCA 214
```

Search completed: March 5, 2006, 22:36:28
Job time : 8.38752 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 23:29:31 ; Search time 43.1372 Seconds
(without alignments)
5750.985 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208
Perfect score: 30
Sequence: 1 catgcttgttcaaccgttcgtcttgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30	100.0	30	8	US-10-713-381-4	Sequence 4, Appli
	2	30	100.0	255	8	US-10-713-381-9	Sequence 9, Appli
	3	30	100.0	1394	8	US-10-713-381-1	Sequence 1, Appli
	4	30	100.0	1394	8	US-10-713-381-2	Sequence 2, Appli
	5	24.2	80.7	158	8	US-10-713-381-3	Sequence 3, Appli
c	6	22	73.3	1304	5	US-10-001-857-90	Sequence 90, Appl
c	7	21	70.0	1980090	8	US-10-719-993-6815	Sequence 6815, Ap
c	8	21	70.0	1980090	8	US-10-741-600-17676	Sequence 17676, A
	9	20.4	68.0	1542	3	US-09-938-842A-1624	Sequence 1624, Ap
	10	20.4	68.0	1542	3	US-09-938-842A-1624	Sequence 1624, Ap
	11	20.2	67.3	2397	7	US-10-282-122A-31883	Sequence 31883, A
	12	20	66.7	2000	7	US-10-260-238-1785	Sequence 1785, Ap
c	13	19.6	65.3	640	7	US-10-437-963-96639	Sequence 96639, A
c	14	19.6	65.3	901	8	US-10-425-115-3534	Sequence 3534, Ap
c	15	19.6	65.3	2000	7	US-10-260-238-1977	Sequence 1977, Ap
	16	19.4	64.7	2243	3	US-09-443-745-15	Sequence 15, Appl
	17	19.4	64.7	14655	10	US-11-063-391-5	Sequence 5, Appli
	18	19.4	64.7	2256646	7	US-10-470-565-1	Sequence 1, Appli
	19	19	63.3	622	4	US-09-925-065A-863989	Sequence 863989,
	20	19	63.3	964	4	US-09-925-065A-546217	Sequence 546217,
c	21	19	63.3	1041	8	US-10-425-115-100318	Sequence 100318,
c	22	19	63.3	1821	7	US-10-425-563-5	Sequence 5, Appli
c	23	19	63.3	1821	9	US-10-972-789A-5	Sequence 5, Appli
c	24	19	63.3	5671	3	US-09-764-872-729	Sequence 729, App
	25	19	63.3	198522	5	US-10-087-192-244	Sequence 244, App
c	26	18.8	62.7	520	4	US-09-925-065A-472473	Sequence 472473,
	27	18.8	62.7	593	4	US-09-925-065A-936359	Sequence 936359,
	28	18.8	62.7	593	4	US-09-925-065A-936360	Sequence 936360,
	29	18.8	62.7	593	4	US-09-925-065A-953397	Sequence 953397,
c	30	18.8	62.7	883	3	US-09-934-455-279	Sequence 279, App
c	31	18.8	62.7	883	6	US-10-225-068-183	Sequence 183, App
c	32	18.8	62.7	883	6	US-10-225-066A-971	Sequence 971, App
c	33	18.8	62.7	883	7	US-10-374-780A-381	Sequence 381, App
c	34	18.8	62.7	883	9	US-10-225-068-183	Sequence 183, App
c	35	18.8	62.7	883	9	US-10-225-066A-971	Sequence 971, App
c	36	18.8	62.7	1083	3	US-09-815-242-7281	Sequence 7281, Ap
c	37	18.8	62.7	1083	7	US-10-282-122A-22639	Sequence 22639, A
c	38	18.8	62.7	1083	7	US-10-335-977-2194	Sequence 2194, Ap
c	39	18.8	62.7	1125	7	US-10-335-977-2195	Sequence 2195, Ap
	40	18.8	62.7	1856	7	US-10-437-963-40048	Sequence 40048, A
c	41	18.8	62.7	1899	8	US-10-472-928-2217	Sequence 2217, Ap
c	42	18.8	62.7	1902	3	US-09-815-242-9220	Sequence 9220, Ap
c	43	18.8	62.7	1902	7	US-10-282-122A-37788	Sequence 37788, A
c	44	18.8	62.7	2000	7	US-10-260-238-2535	Sequence 2535, Ap
c	45	18.8	62.7	2247	6	US-10-369-493-33483	Sequence 33483, A

ALIGNMENTS

RESULT 1
US-10-713-381-4
; Sequence 4, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-4

Query Match 100.0%; Score 30; DB 8; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
||||||||||||||||||||||||||||||
Db 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30

RESULT 2
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME

```
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 255
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-9

  Query Match             100.0%;  Score 30;  DB 8;  Length 255;
  Best Local Similarity   100.0%;  Pred. No. 0.011;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||||||
Db         39 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 68


RESULT 3
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1394
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-1

  Query Match             100.0%;  Score 30;  DB 8;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 0.013;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||||||
Db       1179 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 1208


RESULT 4
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1394
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-2

  Query Match             100.0%;  Score 30;  DB 8;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 0.013;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||||||||||||||||||||||||||
Db       1179 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 1208


RESULT 5
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 158
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-3

  Query Match             80.7%;  Score 24.2;  DB 8;  Length 158;
  Best Local Similarity   89.7%;  Pred. No. 2.8;
  Matches   26;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCC 29
              |||||||||||||||||||||||||  | |
Db         25 CATGCTTGTTCAACCGTTCGTCTTTGTTC 53


RESULT 6
US-10-001-857-90/c
; Sequence 90, Application US/10001857
; Publication No. US20020183500A1
; GENERAL INFORMATION:
;  APPLICANT: Macina, Roberto
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Chen, Sei-Yu
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
;  FILE REFERENCE: DEX-0273
;  CURRENT APPLICATION NUMBER: US/10/001,857
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 60/252,054
;  PRIOR FILING DATE: 2000-11-20
;  NUMBER OF SEQ ID NOS: 208
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 90
;   LENGTH: 1304
;   TYPE: DNA
;   ORGANISM: Homo sapien
```

US-10-001-857-90

Query Match 73.3%; Score 22; DB 5; Length 1304;
Best Local Similarity 83.3%; Pred. No. 31;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
Db 1155 CTTTCTTGTTCCCCCGTTCCTCTTGTTCCA 1126

RESULT 7
US-10-719-993-6815/c
; Sequence 6815, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6815
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6815

Query Match 70.0%; Score 21; DB 8; Length 1980090;
Best Local Similarity 82.8%; Pred. No. 2.2e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
Db 1948918 ATTCTTGCTGAACAGTTCGTCTTGTACCA 1948890

RESULT 8
US-10-741-600-17676/c
; Sequence 17676, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17676
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17676

Query Match 70.0%; Score 21; DB 8; Length 1980090;
Best Local Similarity 82.8%; Pred. No. 2.2e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
Db 1948918 ATTCTTGCTGAACAGTTCGTCTTGTACCA 1948890

RESULT 9
US-09-938-842A-1624
; Sequence 1624, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1624
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1624

Query Match 68.0%; Score 20.4; DB 3; Length 1542;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
Db 237 CATGATTGTTCAACCAATCGTCGGTTTCCA 266

RESULT 10
US-09-938-842A-1624
; Sequence 1624, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME} AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1624
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1624

Query Match 68.0%; Score 20.4; DB 3; Length 1542;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
Db 237 CATGATTGTTCAACCAATCGTCGGTTTCCA 266

```
RESULT 11
US-10-282-122A-31883
; Sequence 31883, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31883
;   LENGTH: 2397
;   TYPE: DNA
;   ORGANISM: Pseudomonas putida
US-10-282-122A-31883

  Query Match             67.3%;  Score 20.2;  DB 7;  Length 2397;
  Best Local Similarity   88.0%;  Pred. No. 2e+02;
  Matches   22;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTG 25
              |||||| || |||||||||| ||||
Db        516 CATGCTGGTACAACCGTTCGGCTTG 540


RESULT 12
US-10-260-238-1785
; Sequence 1785, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
;  APPLICANT: Budworth, Paul R.
;  APPLICANT:  Moughamer, Todd G.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Cooper, Bret
;  APPLICANT:  Glazebrook, Jane
;  APPLICANT:  Goff, Stephen A.
;  APPLICANT:  Katagiri, Fumiyaki
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Provart, Nicholas
;  APPLICANT:  Ricke, Darrell
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
;  FILE REFERENCE: 60111-NP
;  CURRENT APPLICATION NUMBER: US/10/260,238
;  CURRENT FILING DATE:  2002-09-26
;  PRIOR APPLICATION NUMBER: US 60/325,448
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: US 60/325,277
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: US 60/370,620
;  PRIOR FILING DATE: 2002-04-04
;  NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1785
;   LENGTH: 2000
;   TYPE: DNA
;   ORGANISM: Oryza sativa
US-10-260-238-1785

  Query Match             66.7%;  Score 20;  DB 7;  Length 2000;
  Best Local Similarity   82.1%;  Pred. No. 2.3e+02;
  Matches   23;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 TGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              | | |||| |||||||||||||| ||| |
Db       1179 TACATGTTTAACCGTTCGTCTTATTCAA 1206


RESULT 13
US-10-437-963-96639/c
; Sequence 96639, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 96639
;   LENGTH: 640
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_94716C.1
US-10-437-963-96639

  Query Match             65.3%;  Score 19.6;  DB 7;  Length 640;
  Best Local Similarity   84.6%;  Pred. No. 2.9e+02;
  Matches   22;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 CTTGTTCAACCGTTCGTCTTGTTCCA 30
              | |||| |||||||||||||| ||| |
Db        199 CATGTTTAACCGTTCGTCTTATTCAA 174


RESULT 14
US-10-425-115-3534/c
; Sequence 3534, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
```

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 3534
; LENGTH: 901
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_103225C.1
US-10-425-115-3534

Query Match 65.3%; Score 19.6; DB 8; Length 901;
Best Local Similarity 84.6%; Pred. No. 3.1e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 CATGCTTGTTCAACCGTTCGTCTTGT 26
              || |||||||||||||||| |  |||
Db         33 CACGCTTGTTCAACCGTTCCTGGTGT 8
```

RESULT 15
US-10-260-238-1977/c
; Sequence 1977, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1977
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-260-238-1977

Query Match 65.3%; Score 19.6; DB 7; Length 2000;
Best Local Similarity 84.6%; Pred. No. 3.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          5 CTTGTTCAACCGTTCGTCTTGTTCCA 30
              | |||| |||||||||||||| ||| |
Db        291 CATGTTTAACCGTTCGTCTTATTCAA 266
```

Search completed: March 6, 2006, 03:55:17
Job time : 45.1372 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 01:58:03 ; Search time 31.3166 Seconds
(without alignments)
2100.307 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208
Perfect score: 30
Sequence: 1 catgcttgttcaaccgttcgtcttgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		19.4	64.7	2422	12	US-11-136-527-2272	Sequence 2272, Ap
2		19	63.3	622	6	US-09-925-065A-863989	Sequence 863989,
3		19	63.3	964	6	US-09-925-065A-546217	Sequence 546217,
c 4		19	63.3	1821	9	US-11-197-133A-5	Sequence 5, Appli
c 5		18.8	62.7	520	6	US-09-925-065A-472473	Sequence 472473,
6		18.8	62.7	593	6	US-09-925-065A-936359	Sequence 936359,
7		18.8	62.7	593	6	US-09-925-065A-936360	Sequence 936360,
8		18.8	62.7	593	6	US-09-925-065A-953397	Sequence 953397,
c 9		18.8	62.7	1908	7	US-10-523-503-37	Sequence 37, Appl
c 10		18.8	62.7	128963	7	US-10-330-773-86	Sequence 86, Appl
11		18.4	61.3	599	6	US-09-925-065A-811550	Sequence 811550,
12		18.4	61.3	624	6	US-09-925-065A-772237	Sequence 772237,
c 13		18.4	61.3	629	6	US-09-925-065A-787145	Sequence 787145,
14		18.4	61.3	642	6	US-09-925-065A-423	Sequence 423, App
15		18.4	61.3	1012	8	US-10-750-185-52713	Sequence 52713, A
16		18.4	61.3	1012	8	US-10-750-623-52713	Sequence 52713, A
c 17		18.4	61.3	1257	8	US-10-750-185-64410	Sequence 64410, A
c 18		18.4	61.3	1257	8	US-10-750-623-64410	Sequence 64410, A
19		18.4	61.3	1534	6	US-09-925-065A-70688	Sequence 70688, A
20		18.4	61.3	1534	6	US-09-925-065A-70689	Sequence 70689, A
c 21		18	60.0	572	6	US-09-925-065A-819344	Sequence 819344,
22		18	60.0	573	6	US-09-925-065A-821053	Sequence 821053,
23		18	60.0	1395	6	US-09-925-065A-30378	Sequence 30378, A
24		18	60.0	1395	6	US-09-925-065A-30379	Sequence 30379, A
25		17.8	59.3	201	8	US-10-995-561-9774	Sequence 9774, Ap
26		17.8	59.3	201	8	US-10-995-561-9787	Sequence 9787, Ap
27		17.8	59.3	201	8	US-10-995-561-52210	Sequence 52210, A
28		17.8	59.3	603	6	US-09-925-065A-826915	Sequence 826915,
29		17.8	59.3	617	6	US-09-925-065A-673672	Sequence 673672,
30		17.8	59.3	617	6	US-09-925-065A-673673	Sequence 673673,
31		17.8	59.3	652	6	US-09-925-065A-667982	Sequence 667982,
32		17.8	59.3	1385	6	US-09-925-065A-716810	Sequence 716810,
33		17.8	59.3	1907	12	US-11-010-239-68	Sequence 68, Appl
34		17.8	59.3	2039	6	US-09-925-065A-704153	Sequence 704153,
35		17.8	59.3	2039	6	US-09-925-065A-704154	Sequence 704154,
36		17.8	59.3	2039	6	US-09-925-065A-704155	Sequence 704155,
37		17.8	59.3	2867	8	US-10-995-561-357	Sequence 357, App
38		17.8	59.3	3431	8	US-10-995-561-358	Sequence 358, App
39		17.8	59.3	9474	12	US-11-155-232-1	Sequence 1, Appli
40		17.8	59.3	45038	8	US-10-995-561-13311	Sequence 13311, A
41		17.8	59.3	81210	8	US-10-995-561-13295	Sequence 13295, A
c 42		17.8	59.3	163317	12	US-11-117-187-212	Sequence 212, App
43		17.8	59.3	171732	12	US-11-121-086-98	Sequence 98, Appl
44		17.8	59.3	1082144	12	US-11-117-187-211	Sequence 211, App
45		17.6	58.7	600	12	US-11-136-527-7828	Sequence 7828, Ap

ALIGNMENTS

RESULT 1
US-11-136-527-2272
; Sequence 2272, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2272
; LENGTH: 2422
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2272

Query Match 64.7%; Score 19.4; DB 12; Length 2422;
Best Local Similarity 79.3%; Pred. No. 28;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCC 29
              | |||||| ||||| | ||||||||| |
Db          5 CCTGCTTGCTCAACTCTACGTCTTGTTTC 33
```

RESULT 2
US-09-925-065A-863989
; Sequence 863989, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24

```
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 863989
;   LENGTH: 622
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-863989

  Query Match             63.3%;  Score 19;  DB 6;  Length 622;
  Best Local Similarity   81.5%;  Pred. No. 31;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTC 28
              | ||| || ||||| || |||||||||
Db        400 AGGCTAGTCCAACCTTTTGTCTTGTTC 426


RESULT 3
US-09-925-065A-546217
; Sequence 546217, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 546217
;   LENGTH: 964
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-546217

  Query Match             63.3%;  Score 19;  DB 6;  Length 964;
  Best Local Similarity   81.5%;  Pred. No. 35;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTC 28
              | ||| || ||||| || |||||||||
Db        256 AGGCTAGTCCAACCTTTTGTCTTGTTC 282


RESULT 4
US-11-197-133A-5/c
; Sequence 5, Application US/11197133A
; Publication No. US20060040361A1
; GENERAL INFORMATION:
;  APPLICANT: De Le Fuente Jose de Jesus
;  APPLICANT:  Kocan Katherine M.
;  APPLICANT:  Garcia-Almazan Consuelo
;  APPLICANT:  Blouin Edwin F.
;  TITLE OF INVENTION: Protective antigens and vaccines for the control of multi species
;  TITLE OF INVENTION:  infestations
;  FILE REFERENCE: 57338/05-261
;  CURRENT APPLICATION NUMBER: US/11/197,133A
;  CURRENT FILING DATE:  2005-08-04
;  PRIOR APPLICATION NUMBER: US 10/972789
;  PRIOR FILING DATE: 2004-10-25
;  PRIOR APPLICATION NUMBER: US 10/425563
;  PRIOR FILING DATE: 2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/376251
;  PRIOR FILING DATE: 2002-04-29
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
;   LENGTH: 1821
;   TYPE: DNA
;   ORGANISM: Ixodes scapularis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1487)..(1487)
;   OTHER INFORMATION: n is a c g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1595)..(1595)
;   OTHER INFORMATION: n is a c g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1606)..(1606)
;   OTHER INFORMATION: n is a c g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1623)..(1623)
;   OTHER INFORMATION: n is a c g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1762)..(1762)
;   OTHER INFORMATION: n is a c g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1789)..(1789)
;   OTHER INFORMATION: n is a c g or t
US-11-197-133A-5

  Query Match             63.3%;  Score 19;  DB 9;  Length 1821;
  Best Local Similarity   81.5%;  Pred. No. 41;
  Matches   22;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 TGCTTGTTCAACCGTTCGTCTTGTTCC 29
              || || || |||| ||||| |||||||
Db        914 TGGTTCTTGAACCATTCGTGTTGTTCC 888


RESULT 5
US-09-925-065A-472473/c
; Sequence 472473, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 472473
;   LENGTH: 520
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-472473

  Query Match             62.7%;  Score 18.8;  DB 6;  Length 520;
  Best Local Similarity   76.7%;  Pred. No. 37;
  Matches   23;  Conservative     0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||||||| | || |||| ||| ||||||
Db        211 CATGCTTTGTTAAACGTTTCTCTGGTTCCA 182


RESULT 6
US-09-925-065A-936359
; Sequence 936359, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 936359
;   LENGTH: 593
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-936359

  Query Match             62.7%;  Score 18.8;  DB 6;  Length 593;
  Best Local Similarity   76.7%;  Pred. No. 38;
  Matches   23;  Conservative     0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||| |||| ||| ||  |||||||  ||||
Db        341 CATCCTTGCTCACCCTCTCGTCTTTGTCCA 370


RESULT 7
US-09-925-065A-936360
; Sequence 936360, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 936360
;   LENGTH: 593
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-936360

  Query Match             62.7%;  Score 18.8;  DB 6;  Length 593;
  Best Local Similarity   76.7%;  Pred. No. 38;
  Matches   23;  Conservative     0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||| |||| ||| ||  |||||||  ||||
Db        341 CATCCTTGCTCACCCTCTCGTCTTTGTCCA 370


RESULT 8
US-09-925-065A-953397
; Sequence 953397, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 953397
;   LENGTH: 593
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-953397

  Query Match             62.7%;  Score 18.8;  DB 6;  Length 593;
  Best Local Similarity   76.7%;  Pred. No. 38;
  Matches   23;  Conservative     0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              ||| |||| ||| ||  |||||||  ||||
Db        341 CATCCTTGCTCACCCTCTCGTCTTTGTCCA 370


RESULT 9
US-10-523-503-37/c
; Sequence 37, Application US/10523503
; Publication No. US20060037102A1
; GENERAL INFORMATION:
;  APPLICANT: BASF PLANT SCIENCE GMBH
;  TITLE OF INVENTION: SUGAR AND LIPID METABOLISM REGULATORS IN PLANTS IV
;  FILE REFERENCE: 16313-0236
;  CURRENT APPLICATION NUMBER: US/10/523,503
;  CURRENT FILING DATE:  2005-02-02
;  PRIOR APPLICATION NUMBER: US 60/400,803
;  PRIOR FILING DATE: 2002-08-02
;  NUMBER OF SEQ ID NOS: 163
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
;   LENGTH: 1908
```

```
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-10-523-503-37

  Query Match             62.7%;  Score 18.8;  DB 7;  Length 1908;
  Best Local Similarity   76.7%;  Pred. No. 51;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              | |||||| ||| ||||| | | ||| ||||
Db       1688 CTTGCTTCTTCGACCGTGCCTTTTGGTCCA 1659


RESULT 10
US-10-330-773-86/c
; Sequence 86, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 86
;   LENGTH: 128963
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(128963)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-86

  Query Match             62.7%;  Score 18.8;  DB 7;  Length 128963;
  Best Local Similarity   76.7%;  Pred. No. 1.5e+02;
  Matches   23;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 CATGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              || | || |||||||||||| || |||  ||
Db     102327 CAAGTTTCTTCAACCGTTCTTCATGTCACA 102298


RESULT 11
US-09-925-065A-811550
; Sequence 811550, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 811550
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-811550

  Query Match             61.3%;  Score 18.4;  DB 6;  Length 599;
  Best Local Similarity   78.6%;  Pred. No. 59;
  Matches   22;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 TGCTTGTTCAACCGTTCGTCTTGTTCCA 30
              |||||| |||||| | |  ||| |||||
Db        504 TGCTTGCTCAACCCTGCTGCTTCTTCCA 531


RESULT 12
US-09-925-065A-772237
; Sequence 772237, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 772237
;   LENGTH: 624
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-772237

  Query Match             61.3%;  Score 18.4;  DB 6;  Length 624;
  Best Local Similarity   78.6%;  Pred. No. 60;
  Matches   22;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 ATGCTTGTTCAACCGTTCGTCTTGTTCC 29
              ||  ||||||||| ||||||  |||| |
Db         49 ATTTTTGTTCAACTGTTCGTTCTGTTTC 76


RESULT 13
US-09-925-065A-787145/c
; Sequence 787145, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 787145
;   LENGTH: 629
```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-787145

  Query Match 61.3%; Score 18.4; DB 6; Length 629;
  Best Local Similarity 78.6%; Pred. No. 60;
  Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ATGCTTGTTCAACCGTTCGTCTTGTTCC 29
  || |||||||| |||||| |||| |
Db 576 ATTTTTGTTCAACTGTTCGTTCTGTTTC 549

RESULT 14
US-09-925-065A-423
; Sequence 423, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 423
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-423

  Query Match 61.3%; Score 18.4; DB 6; Length 642;
  Best Local Similarity 78.6%; Pred. No. 60;
  Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 TGCTTGTTCAACCGTTCGTCTTGTTCCA 30
  |||||| || || |||||||| | |||
Db 295 TGCTTGCTCCTCCATTCGTCTTTTGCCA 322

RESULT 15
US-10-750-185-52713
; Sequence 52713, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 52713
; LENGTH: 1012
; TYPE: DNA
; ORGANISM: Bovine 19866881292329
US-10-750-185-52713

  Query Match 61.3%; Score 18.4; DB 8; Length 1012;
  Best Local Similarity 78.6%; Pred. No. 67;
  Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ATGCTTGTTCAACCGTTCGTCTTGTTCC 29
  || |||||| |||||| ||||| |||
Db 454 ATTCTTGTTTAACCGTATTTCTTGCTCC 481

Search completed: March 6, 2006, 10:32:51
Job time : 31.3166 secs

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:38:03 ; Search time 174.564 Seconds
(without alignments)
13025.213 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278
Perfect score: 40
Sequence: 1 aggatacctactcccaaacaatccatcttactcatgcaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	100.0	40	6	AX224399	AX224399 Sequence
2	40	100.0	50	6	AX224398	AX224398 Sequence
3	40	100.0	158	6	AX224396	AX224396 Sequence
4	40	100.0	255	6	AX224402	AX224402 Sequence
5	40	100.0	1394	6	BD062176	BD062176 Male tiss
6	40	100.0	1394	6	BD062177	BD062177 Male tiss
7	40	100.0	1394	6	AX224394	AX224394 Sequence
8	40	100.0	1394	6	AX224395	AX224395 Sequence
9	40	100.0	3343	15	AF360356	AF360356 Zea mays
c 10	30.4	76.0	110000	15	AP008209_086	Continuation (87 o
c 11	30.4	76.0	137327	15	AC135206	AC135206 Oryza sat
12	25.2	63.0	303308	14	AC156670	AC156670 Bos tauru
c 13	24	60.0	55001	8	AC084290	AC084290 Homo sapi
c 14	24	60.0	127196	9	AL807784	AL807784 Mouse DNA
15	24	60.0	154857	8	AC157528	AC157528 Pan trogl
c 16	24	60.0	171940	14	AC148514	AC148514 Papio anu
17	24	60.0	216548	14	AC161620	AC161620 Pan trogl
c 18	23.6	59.0	193735	14	AC148501	AC148501 Callithri
c 19	23.2	58.0	170392	14	AC156784	AC156784 Rhinoloph
c 20	23.2	58.0	182616	14	AC160916	AC160916 Rhinoloph
21	23.2	58.0	211584	9	AC153927	AC153927 Mus muscu
c 22	23	57.5	52648	5	CR788256	CR788256 Zebrafish
23	23	57.5	117026	8	AC008897	AC008897 Homo sapi
c 24	23	57.5	153971	9	AC102478	AC102478 Mus muscu
25	23	57.5	177104	14	AC125876	AC125876 Rattus no
c 26	23	57.5	195285	14	CR847932	CR847932 Danio rer
27	23	57.5	218224	9	AC134404	AC134404 Mus muscu
c 28	23	57.5	234163	14	AC115318	AC115318 Rattus no
29	23	57.5	235286	14	AC159385	AC159385 Bos tauru
30	23	57.5	242109	14	AC147523	AC147523 Otolemur
c 31	23	57.5	319056	14	AC016640	AC016640 Homo sapi
32	22.8	57.0	164944	2	AC008195	AC008195 Drosophil
33	22.8	57.0	175781	2	AC008091	AC008091 Drosophil
34	22.8	57.0	235928	2	AE003736	AE003736 Drosophil
35	22.8	57.0	309657	14	AC157205	AC157205 Bos tauru
c 36	22.6	56.5	90463	14	AC165458	AC165458 Bos tauru
37	22.6	56.5	95896	8	AL356384	AL356384 Human DNA
38	22.6	56.5	129410	14	AC149638	AC149638 Medicago
39	22.6	56.5	160246	14	AC009221	AC009221 Homo sapi
40	22.6	56.5	161817	9	AC132119	AC132119 Mus muscu
41	22.6	56.5	170973	8	AP000923	AP000923 Homo sapi
42	22.6	56.5	175110	8	AP002428	AP002428 Homo sapi
43	22.6	56.5	186457	14	AC129669	AC129669 Rattus no
44	22.6	56.5	196990	14	AC134013	AC134013 Rattus no
c 45	22.6	56.5	235953	14	AC130117	AC130117 Rattus no

ALIGNMENTS

RESULT 1
AX224399
LOCUS AX224399 40 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 6 from Patent WO0160997.
ACCESSION AX224399
VERSION AX224399.1 GI:15554641
KEYWORDS .
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 6 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES Location/Qualifiers
source 1. .40
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
||||||||||||||||||||||||||||||||||||||||
Db 1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40

RESULT 2
AX224398
LOCUS AX224398 50 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 5 from Patent WO0160997.
ACCESSION AX224398
VERSION AX224398.1 GI:15554640
KEYWORDS .
SOURCE Zea mays

```
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 5 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             100.0%;  Score 40;  DB 6;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 0.0001;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db         11 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 50


RESULT 3
AX224396
LOCUS       AX224396                158 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 3 from Patent WO0160997.
ACCESSION   AX224396
VERSION     AX224396.1  GI:15554638
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 3 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .158
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             100.0%;  Score 40;  DB 6;  Length 158;
  Best Local Similarity   100.0%;  Pred. No. 8.6e-05;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db         86 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 125


RESULT 4
AX224402
LOCUS       AX224402                255 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 9 from Patent WO0160997.
ACCESSION   AX224402
VERSION     AX224402.1  GI:15554644
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 9 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .255
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             100.0%;  Score 40;  DB 6;  Length 255;
  Best Local Similarity   100.0%;  Pred. No. 8e-05;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db         99 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 138


RESULT 5
BD062176
LOCUS       BD062176                1394 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Male tissue-preferred regulatory region and method of using same.
ACCESSION   BD062176
VERSION     BD062176.1  GI:22607781
KEYWORDS    JP 2001520523-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1394)
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
            Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: JP 2001520523-A 1 30-OCT-2001;
            PIONEER HI BRED INTERNATIONAL INC
COMMENT     PN   JP 2001520523-A/1
            PD   30-OCT-2001
            PF   19-JUN-1998 JP 1999504910
            PR   23-JUN-1997 US    08/880499
            PI   MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
            PI   TIMMY L KENDALL
            PC   C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
            ,C07K14/34,C12Q1/68,
            PC   A01H5/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .1394
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 40;  DB 6;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-05;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 6
BD062177
LOCUS       BD062177                1394 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Male tissue-preferred regulatory region and method of using same.
ACCESSION   BD062177
VERSION     BD062177.1  GI:22607782
KEYWORDS    JP 2001520523-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1394)
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
            Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: JP 2001520523-A 2 30-OCT-2001;
            PIONEER HI BRED INTERNATIONAL INC
COMMENT     PN   JP 2001520523-A/2
            PD   30-OCT-2001
            PF   19-JUN-1998 JP 1999504910
            PR   23-JUN-1997 US    08/880499
            PI   MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
            PI   TIMMY L KENDALL
            PC   C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
            ,C07K14/34,C12Q1/68,
            PC   A01H5/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES            Location/Qualifiers
     source         1. .1394
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
ORIGIN

  Query Match           100.0%;  Score 40;  DB 6;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-05;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 7
AX224394
LOCUS       AX224394               1394 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 1 from Patent WO0160997.
ACCESSION   AX224394
VERSION     AX224394.1  GI:15554636
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 1 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES            Location/Qualifiers
     source         1. .1394
                    /organism="Zea mays"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:4577"
ORIGIN

  Query Match           100.0%;  Score 40;  DB 6;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-05;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 8
AX224395
LOCUS       AX224395               1394 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 2 from Patent WO0160997.
ACCESSION   AX224395
VERSION     AX224395.1  GI:15554637
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 2 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES            Location/Qualifiers
     source         1. .1394
                    /organism="Zea mays"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:4577"
ORIGIN

  Query Match           100.0%;  Score 40;  DB 6;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-05;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 9
AF360356
LOCUS       AF360356               3343 bp    DNA     linear   PLN 12-MAY-2001
DEFINITION  Zea mays male fertility protein (Ms45) gene, complete cds.
ACCESSION   AF360356
VERSION     AF360356.1  GI:14028756
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3343)
  AUTHORS   Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
  TITLE     Cloning of Ms45, a gene required for male ferility from Zea mays
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3343)
  AUTHORS   Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
            Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
            IA 50131-1004, USA
FEATURES            Location/Qualifiers
     source         1..3343
                    /organism="Zea mays"
                    /mol_type="genomic DNA"
                    /cultivar="B73"
                    /db_xref="taxon:4577"
                    /chromosome="9L"
     gene           <1392. .>2942
                    /gene="Ms45"
     mRNA           join(<1392. .1768,1898. .2182,2280. .2447,2534. .>2942)
                    /gene="Ms45"
                    /product="male fertility protein"
     CDS            join(1392. .1768,1898. .2182,2280. .2447,2534. .2942)
                    /gene="Ms45"
                    /codon_start=1
                    /product="male fertility protein"
                    /protein_id="AAK52489.1"
                    /db_xref="GI:14028757"
                    /translation="MEKRNLQWRRGRDGIVQYPHLFFAALALALLVADPPGLSPLAEV
                    DYRPVKHELAPYGEVMGSWPRDNASRLRRGRLEFVGEVFGPESIEFDLQGRGPYAGLA
                    DGRVVRWMGEEAGWETFAVMNPDWSEEVCANGVNSTTRKQHEKEEFCGRPLGLRFHGE
```

TGELYVADAYYGLMVVGQSGGVASSVAREADGDPIRFANDLDVHRNGSVFFTDTSMRY
SRKDHLNILLEGEGTGRLLRYDPETSGVHVVLKGLVFPNGVQISEDHQFLLFSETTNC
RIMRYWLEGPRAGEVEVFANLPGFPDNVRSNGRGQFWVAIDCCRTPAQEVFAKRPWLR
TLYFKFPLSLKVLTWKAARRMHTVLALLDGEGRVVEVLEDRGHEVMKLVSEVREVGRK
LWIGTVAHNHIATIPYPLED"

ORIGIN

Query Match 100.0%; Score 40; DB 15; Length 3343;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
||||||||||||||||||||||||||||||||||||||||
Db 1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278

RESULT 10
AP008209_086/c
WPCOMMENT
Sequence split into 362 fragments LOCUS AP008209 Accession AP008209

Fragment Name	Begin	End
AP008209_000	1	110000
AP008209_001	100001	210000
AP008209_002	200001	310000
AP008209_003	300001	410000
AP008209_004	400001	510000
AP008209_005	500001	610000
AP008209_006	600001	710000
AP008209_007	700001	810000
AP008209_008	800001	910000
AP008209_009	900001	1010000
AP008209_010	1000001	1110000
AP008209_011	1100001	1210000
AP008209_012	1200001	1310000
AP008209_013	1300001	1410000
AP008209_014	1400001	1510000
AP008209_015	1500001	1610000
AP008209_016	1600001	1710000
AP008209_017	1700001	1810000
AP008209_018	1800001	1910000
AP008209_019	1900001	2010000
AP008209_020	2000001	2110000
AP008209_021	2100001	2210000
AP008209_022	2200001	2310000
AP008209_023	2300001	2410000
AP008209_024	2400001	2510000
AP008209_025	2500001	2610000
AP008209_026	2600001	2710000
AP008209_027	2700001	2810000
AP008209_028	2800001	2910000
AP008209_029	2900001	3010000
AP008209_030	3000001	3110000
AP008209_031	3100001	3210000
AP008209_032	3200001	3310000
AP008209_033	3300001	3410000
AP008209_034	3400001	3510000
AP008209_035	3500001	3610000
AP008209_036	3600001	3710000
AP008209_037	3700001	3810000
AP008209_038	3800001	3910000
AP008209_039	3900001	4010000
AP008209_040	4000001	4110000
AP008209_041	4100001	4210000
AP008209_042	4200001	4310000
AP008209_043	4300001	4410000
AP008209_044	4400001	4510000
AP008209_045	4500001	4610000
AP008209_046	4600001	4710000
AP008209_047	4700001	4810000
AP008209_048	4800001	4910000
AP008209_049	4900001	5010000
AP008209_050	5000001	5110000
AP008209_051	5100001	5210000
AP008209_052	5200001	5310000
AP008209_053	5300001	5410000
AP008209_054	5400001	5510000
AP008209_055	5500001	5610000
AP008209_056	5600001	5710000
AP008209_057	5700001	5810000
AP008209_058	5800001	5910000
AP008209_059	5900001	6010000
AP008209_060	6000001	6110000
AP008209_061	6100001	6210000
AP008209_062	6200001	6310000
AP008209_063	6300001	6410000
AP008209_064	6400001	6510000
AP008209_065	6500001	6610000
AP008209_066	6600001	6710000
AP008209_067	6700001	6810000
AP008209_068	6800001	6910000
AP008209_069	6900001	7010000
AP008209_070	7000001	7110000
AP008209_071	7100001	7210000
AP008209_072	7200001	7310000
AP008209_073	7300001	7410000
AP008209_074	7400001	7510000
AP008209_075	7500001	7610000
AP008209_076	7600001	7710000
AP008209_077	7700001	7810000
AP008209_078	7800001	7910000
AP008209_079	7900001	8010000
AP008209_080	8000001	8110000
AP008209_081	8100001	8210000
AP008209_082	8200001	8310000
AP008209_083	8300001	8410000
AP008209_084	8400001	8510000
AP008209_085	8500001	8610000
AP008209_086	8600001	8710000
AP008209_087	8700001	8810000
AP008209_088	8800001	8910000
AP008209_089	8900001	9010000
AP008209_090	9000001	9110000
AP008209_091	9100001	9210000
AP008209_092	9200001	9310000
AP008209_093	9300001	9410000
AP008209_094	9400001	9510000
AP008209_095	9500001	9610000
AP008209_096	9600001	9710000
AP008209_097	9700001	9810000
AP008209_098	9800001	9910000
AP008209_099	9900001	10010000
AP008209_100	10000001	10110000
AP008209_101	10100001	10210000
AP008209_102	10200001	10310000
AP008209_103	10300001	10410000
AP008209_104	10400001	10510000
AP008209_105	10500001	10610000
AP008209_106	10600001	10710000
AP008209_107	10700001	10810000
AP008209_108	10800001	10910000
AP008209_109	10900001	11010000
AP008209_110	11000001	11110000
AP008209_111	11100001	11210000
AP008209_112	11200001	11310000
AP008209_113	11300001	11410000
AP008209_114	11400001	11510000
AP008209_115	11500001	11610000
AP008209_116	11600001	11710000
AP008209_117	11700001	11810000
AP008209_118	11800001	11910000
AP008209_119	11900001	12010000
AP008209_120	12000001	12110000
AP008209_121	12100001	12210000
AP008209_122	12200001	12310000
AP008209_123	12300001	12410000
AP008209_124	12400001	12510000

AP008209_125	12500001	12610000
AP008209_126	12600001	12710000
AP008209_127	12700001	12810000
AP008209_128	12800001	12910000
AP008209_129	12900001	13010000
AP008209_130	13000001	13110000
AP008209_131	13100001	13210000
AP008209_132	13200001	13310000
AP008209_133	13300001	13410000
AP008209_134	13400001	13510000
AP008209_135	13500001	13610000
AP008209_136	13600001	13710000
AP008209_137	13700001	13810000
AP008209_138	13800001	13910000
AP008209_139	13900001	14010000
AP008209_140	14000001	14110000
AP008209_141	14100001	14210000
AP008209_142	14200001	14310000
AP008209_143	14300001	14410000
AP008209_144	14400001	14510000
AP008209_145	14500001	14610000
AP008209_146	14600001	14710000
AP008209_147	14700001	14810000
AP008209_148	14800001	14910000
AP008209_149	14900001	15010000
AP008209_150	15000001	15110000
AP008209_151	15100001	15210000
AP008209_152	15200001	15310000
AP008209_153	15300001	15410000
AP008209_154	15400001	15510000
AP008209_155	15500001	15610000
AP008209_156	15600001	15710000
AP008209_157	15700001	15810000
AP008209_158	15800001	15910000
AP008209_159	15900001	16010000
AP008209_160	16000001	16110000
AP008209_161	16100001	16210000
AP008209_162	16200001	16310000
AP008209_163	16300001	16410000
AP008209_164	16400001	16510000
AP008209_165	16500001	16610000
AP008209_166	16600001	16710000
AP008209_167	16700001	16810000
AP008209_168	16800001	16910000
AP008209_169	16900001	17010000
AP008209_170	17000001	17110000
AP008209_171	17100001	17210000
AP008209_172	17200001	17310000
AP008209_173	17300001	17410000
AP008209_174	17400001	17510000
AP008209_175	17500001	17610000
AP008209_176	17600001	17710000
AP008209_177	17700001	17810000
AP008209_178	17800001	17910000
AP008209_179	17900001	18010000
AP008209_180	18000001	18110000
AP008209_181	18100001	18210000
AP008209_182	18200001	18310000
AP008209_183	18300001	18410000
AP008209_184	18400001	18510000
AP008209_185	18500001	18610000
AP008209_186	18600001	18710000
AP008209_187	18700001	18810000
AP008209_188	18800001	18910000
AP008209_189	18900001	19010000
AP008209_190	19000001	19110000
AP008209_191	19100001	19210000
AP008209_192	19200001	19310000
AP008209_193	19300001	19410000
AP008209_194	19400001	19510000
AP008209_195	19500001	19610000
AP008209_196	19600001	19710000

Query Match 76.0%; Score 30.4; DB 15; Length 110000;
Best Local Similarity 85.0%; Pred. No. 0.18;
Matches 34; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              || | ||||||||||||||||||    || ||||||||||
Db      36341 AGAACACCTACTCCCAAACAATCACGCTGACTCATGCAAC 36302
```

RESULT 11
AC135206/c
LOCUS AC135206 137327 bp DNA linear PLN 16-APR-2003
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 3 clone OJ1041F02, complete sequence.
ACCESSION AC135206
VERSION AC135206.3 GI:27596977
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 137327)
AUTHORS Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and Collura,K.
TITLE Rice Genomic Sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 137327)
AUTHORS Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C., Currie,J., Collura,K. and Thompson,S.
TITLE Direct Submission
JOURNAL Submitted (09-OCT-2002) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE 3 (bases 1 to 137327)
AUTHORS Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C., Currie,J., Collura,K. and Thompson,S.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-2002) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE 4 (bases 1 to 137327)
AUTHORS Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J., Collura,K. and Thompson,S.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-2003) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE 5 (bases 1 to 137327)
AUTHORS Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T., Saski,C., Henry,D., Thompson,S., Simmons,J., Thurmond,S.K. and Sun,S.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2003) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
COMMENT On Jan 11, 2003 this sequence version replaced gi:24635891.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by more than one plasmid subclone ; and the assembly was confirmed by restriction digest.There are transposons located at area 29427-32874 and at area 19945-33212. Region 70511-70536 is covered by a single subclone. Region 18678-18873 is a double stranded single subclone. Areas 69957-70761, 89836-90416, 98951-99009, 105316-105607, and 133052-133152 are covered by Monsanto only. The nucleotide sequence of this BAC clone was generated by combining Syngenta, Monsanto and Arizona Genomics Institute sequencing data.
FEATURES Location/Qualifiers
source 1. .137327
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39947"
/chromosome="3"
/clone="OJ1041F02"

```
                     /note="(japonica cultivar-group)"
     gene            complement(638. .9495)
                     /gene="OJ1041F02.1"
                     /note="unknown protein"
     CDS             complement(join(638. .838,1141. .1276,3132. .3244,3326. .3421,
                     3513. .3727,3812. .3934,4024. .4054,4157. .4216,5088. .5243,
                     5322. .5411,5484. .5699,5808. .5942,6035. .6538,6622. .6837,
                     7548. .7593,7679. .7815,8196. .8328,9221. .9495))
                     /gene="OJ1041F02.1"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="AAP06844.1"
                     /db_xref="GI:29893590"
                     /translation="MAQSLELLLIQFLMPDNDARRQAEEQIRRLARDPQVVPALVHHL
                     RTAKTPNVRQLAAVLLRKKITSHWPKLPPHAKASLKQALIDSITIDHSHLVRRASANV
                     VSIIAKYAVPAGEWPELLPFIFQCSQSPQEDHREVALILFSSLTETIGTTFQSHLNDL
                     QPILLKCLQDEASSRVRIAALKAVGSFIEYVNDGGDVVKIFRDFVPSILNVSRQCLAN
                     GEEDVASIAFEIFDELIESPAPLLGDSVRSIVQFSLEVCSNQELEINIRQQAIQIISW
                     LVKFKASFLKKHKLVIPILQVMCPLLTETADEDGDSDLAADRSAAEVIDTMAINLPRH
                     VFPPVLEFASVSFRHINPKYREAAVTSLGVVSEGCCEHLKDKLEDCLKVVLEALKDQE
                     QMVRGAASFALGQFAEHLQPEILSHYESVLPCILNALEDPSDEVKEKSYYALAAFCED
                     MGENILPYLDPLMCRLVMSLQGSPRNLQETCMSAIGSVAAAAEQAFMPYAEKVLEMMK
                     GFMVLTNDEDLCARARATEVVGIVAMAVGRARMETILPPFIEAAISGFVLDYSELREY
                     THGFFSNVAEILDDSFAQYLPHVVPLAFSSCNLDDGSAVDIDDADSVDNGFSGVSSDD
                     DVNDEPRVRNISWLTIIADILTAIRAIPPAHADVLEKQKDILDTVMNIYIKTMREDDD
                     KEVVAQACTSLADIVRDCGFAIIEPYITRLADATLILLRQESCCQQVESDGEDDGDID
                     HDEVLMDAVSDLLPAFAKVMGSYFDPIFTKLFDSLMKFAKSPHPPQDKTMVVATLAEV
                     AQGMGAPISAYVDKIMPLVLKELASSEATNRRNAAFCVGEMCKNGGAAALKYYGDILH
                     GLHRLFADSEPDDAVRDNAAGAIARMIMVQPQSIPLNQILPLVPDVINAFAQVVVSPN
                     ESDEVKTVVAKAVSHLISVYGQQMQPILSALPPAHANALASFANRR"
     gene            12372. .15503
                     /gene="OJ1041F02.2"
                     /note="unknown protein"
     CDS             join(12372. .12757,12858. .13025,13130. .13295,13716. .13826,
                     13982. .14130,14224. .14354,14459. .14558,15197. .15503)
                     /gene="OJ1041F02.2"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="AAP06845.1"
                     /db_xref="GI:29893591"
                     /translation="MPPQQGECCRHGWLGVGEAAVGGGGEEPFYVPLRKRLSVDGKAS
                     TAPRICIWECDGEAGDITCDIVAAPLRRSCSAKAMPPPAPLFRMMTPPPPRPQRGDGE
                     EARRPGEAIRKGHRSYSLMLNLQLGISYSVGKSSALPFQKLAASDFDPREKVWTRFPP
                     EGSKFTPPHHSVDFRWKDYCPAVFRHLRKLFGVDPAEYMLAICGNDTLRELASPGKSG
                     SCFFITQDDRFMIKTVKKSEVKVLIRMLRSYYEHVRQYKSTLLTRFYGTHCIKQAGCP
                     KIHRRFDLKGSSHGRTIDKTERKIDETTTLKDLDLQYAFRLQRFWYEELMKQIQMDCT
                     FLETQGIMDYSLLLGVHFRNDYSVSKIGISQHIAFPKSTGKRKSFEGGSSFCELCFVE
                     SGCKDRDLIDSRKPFIQLGINMPAQAERSSKKILDNFLLNERHLFITPPSGGSCDVYL
                     FFGIIDILQDYDITKKLEHAYKSFQVNPDYISAVDPKLYSRRFQDFIRRVFIKEQ"
     repeat_region   16734. .16878
                     /note="putative MITE, Gaijin/Gaigin-like"
     repeat_region   19263. .19454
                     /note="putative transposon, CACTA, En/Spm-like"
     gene            21819. .24660
                     /gene="OJ1041F02.3"
                     /note="hypothetical protein"
     CDS             join(21819. .22007,22096. .22206,23772. .23996,24523. .24660)
                     /gene="OJ1041F02.3"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAP06846.1"
                     /db_xref="GI:29893592"
                     /translation="MATASDGSGSGGEQRRLLSIPKEGERIIAPTRRPDGTLRKAIRI
                     RAGYVPQEEVAIYQSKGAQMRKSGPDVPPGYDPALDAKPKTKAAKRNERRKEKRQQAS
                     TTNDKGKGLHIEDDAGETDNPKDAVDSVTKQISGIAISESLVVATSSTDATDNSKSES
                     SAPDIDKKIRALKKKIRLAEAQVQGDPENLKPEQLEKMKKIEGWKEELKLLENKSSPA
                     AS"
     gene            complement(26563. .30575)
                     /gene="OJ1041F02.4"
                     /note="unknown protein"
     CDS             complement(join(26563. .26772,27082. .27288,27430. .27537,
                     27881. .27940,28051. .28126,28327. .28484,28616. .28678,
                     28790. .28874,29560. .30575))
                     /gene="OJ1041F02.4"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="AAP06847.1"
                     /db_xref="GI:29893593"
                     /translation="MGPPAAPPSPSSSSGGSSRRRRRLERRNAAKHIGYDASNFCAYP
                     QSPPAASAPASGSPSLACSPACSLDLTSFRIGGSGDGCRDVQLLCSSLGLSGVDDFAV
                     PVADWEAHKAGRSSFSSSASTPKPREEPPARDSPVRREVAAEEEPPSLPAPAAAPVLP
                     AKETPRSVAIEAPAPLLRVDPWEPARPDVRKASGEGGIKGVRPPPVVLKPPPSMVRPA
                     VCVVESTWDILRSFAPEEDSHAHAPASRSGGDSACQDAGEEEDDAAAVLTLEELRLGE
                     TSEEFTGTSSLSTTNDDETSSTTTESMFYISPNGRFRRKIRSWNRGMLLGSGSFGTVF
                     EGISDEGVFFAVKEVCLCDQGSNAQQCIFQLEQEIALLSQFEHENIVQYYGTDKEDSK
                     LYIFLELVTQGSLASLYQKYRLRDTHVSAYTRQILNGLTYLHERNIVHRDIKCANILV
                     HANGSVKLADFGLAKEITKFNVLKSCKGTVYWMAPEVVNPKTTYGPEADIWSLGCTVL
                     EMLTRQLPYPGLEWTQALYRIGKGEPPAIPNCLSRDARDFISQCVKPNPQDRPSAAKL
                     LEHPFVNRSMRSIRSMRTSSRSNSSILVLLIMWSWPGRTIRYREAILSTAQSSTPTIA
                     FHLLILSHSLLRGDMSWFSQRLIFSPENRLNVHSSA"
     repeat_region   27679. .27832
                     /note="putative MITE, MITE-adh, type D-like"
     repeat_region   complement(31668. .31815)
     gene            complement(35049. .49855)
                     /gene="OJ1041F02.5"
                     /note="hypothetical protein"
     CDS             complement(join(35049. .35184,35866. .36033,36922. .37205,
                     37280. .38523,40038. .40104,40196. .40287,40406. .40617,
                     41678. .41760,42229. .42259,42907. .43088,43694. .43753,
                     44229. .44426,44622. .44676,45877. .45992,46149. .46313,
                     46447. .46620,47271. .47470,47915. .48019,48200. .48263,
                     49589. .49855))
                     /gene="OJ1041F02.5"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAP06848.1"
                     /db_xref="GI:29893594"
                     /translation="MGGSADPEAPTPTPSPSPSPAKATPSPASADGNRLRRCVQSKLS
                     WGPPKAGGGGGEAGGAGLPPLAAGDGTPEKVKKRGRPRKSEAGKKPSSNRETTGLEQD
                     SKDEVILVDESPQKKQRKGRGKNQGAALKVPNRKHCKALESTDGHESCQQLRSSQTQA
                     VLPQKSPTSVDIDLVTGPSEASPVNDNVDALDNEDKPQLIVDLRSEANIAAEENRRLS
                     SGKKMHPFFASRKIHKGAGQDILNVEDEDMDSLCAFERDPPLCPVHVLYELEVTMPIH
                     WSNKWLIADKSFLGTSTTEQNSAEHADPGKHLANFHDKQNKSKFSSQDVIDVDDECLL
                     ASSSCFHASLFESKQHERVQHELPEVTPKGCQTANLWTDKYRPETAAQVCGNSKHVKF
                     LNEWLKGWDERGHRNKQNIVTNGSMNGRSCQDGSDTDYSEDASDYENVLLITGPVGCG
                     KSAAVFACAREQGFNVIELNMPKTVVFPFLPHNKLLAQTLDLKPDKLHDSSHPSIKYI
                     PTRVFVHMNRNSIISISAHSSSLSVIKLHCLVNTSDMRNGAYVRQKFEEATKSHGLEK
                     WSQEEIIGLPISNSLDPASGTPGTAEYKQVINKTLILFEDVDTVFDEDRGFISTILKM
                     VETTKWPIILTSNKKDPPLPHLLAQLVLDFTYPSSAELLSHVDMICKSEGVEITVPQQ
                     KHIIDAFLGRLNKCLSCPSLLDLDAVHSTVPRIMPWDFPCKLSETIYMEIDKTIVTAE
                     QKKKQMEVSEFEGLELQIMTPLTKGRSAGKTRKPKKSKLKHGRSADCNDASPCKNDLD
                     DFHDSPDIPLPSNHQRMRNRRGVVLFAESDDDLADAHAAKDATFTVQEGRLLPQSSEL
                     PCLYGHGISNIVPESVFFQQSSVPHLHREVISNQLCFPSESRAFEPASSFQNQLESNM
                     PGSISQICDTFMSQGISCVPESSFMVGGTSASISSDDLLSSLVSNGLSALRNESTYTA
                     SVVALEDTNKVENQMTDKPQKCMEDEVGETCEAYVELADRNDHASCSITGYQLMDECS
                     RAESVWLLSGKKNNDSCKVEHVQDTWNRLRQCHPVLPCDMNHNRSVSGALKRVSRVSD
                     LISESDLMLISCHPFSNDISDPSLTPYTESDGFSYSKQLEMGSIYAQHGLCIFLQDSQ
                     ATDDGFVDFLQELLFSGTTTTSLGKFVSSGISCGDGSGNISHVKYPTSCISKRRERQA
                     RLREVLLPVVPPKLSQSLRGPAFVDYLSSMSQISQLENMQLSECKASSKQRRCRQPRH
                     YLSSGALSLSAEDIGLLAQCSTFSDRRESETIIEQAIS"
     repeat_region   39012. .39207
                     /note="putative MITE, MITE-adh, type M-like"
     repeat_region   39681. .39879
                     /note="putative MITE, MITE-adh, type G-like"
     repeat_region   51474. .51836
                     /note="putative MITE, Castaway-like"
     gene            complement(59793. .62196)
                     /gene="OJ1041F02.6"

  Query Match           76.0%;  Score 30.4;  DB 15;  Length 137327;
  Best Local Similarity  85.0%;  Pred. No. 0.17;
  Matches   34;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              || | ||||||||||||||||||   || ||||||||||
Db     119478 AGAACACCTACTCCCAAACAATCACGCTGACTCATGCAAC 119439
```

```
RESULT 12
AC156670
LOCUS       AC156670               303308 bp    DNA     linear   HTG 01-JUL-2005
DEFINITION  Bos taurus clone CH240-60K24, *** SEQUENCING IN PROGRESS ***, 48
            unordered pieces.
ACCESSION   AC156670
VERSION     AC156670.2  GI:68266432
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 303308)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,P., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 303308)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 303308)
  AUTHORS   Cow Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jun 28, 2005 this sequence version replaced gi:58531390.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: FDJU
                Center clone name: CH240-60K24
            -------------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 284059 bases at least Q40
                Consensus quality: 288397 bases at least Q30
                Consensus quality: 292413 bases at least Q20
                Estimated insert size: 289468; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
            * NOTE: This sequence may represent more than one clone.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 48 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     4448: contig of 4448 bp in length
            *     4449     4498: gap of 50 bp
            *     4499    14243: contig of 9745 bp in length
            *    14244    14293: gap of 50 bp
            *    14294    35369: contig of 21076 bp in length
            *    35370    35419: gap of 50 bp
            *    35420    47505: contig of 12086 bp in length
            *    47506    47555: gap of 50 bp
            *    47556    49432: contig of 1877 bp in length
            *    49433    49482: gap of 50 bp
            *    49483    51615: contig of 2133 bp in length
            *    51616    52521: gap of 906 bp
            *    52522    59438: contig of 6917 bp in length
            *    59439    59488: gap of 50 bp
            *    59489    64821: contig of 5333 bp in length
            *    64822    64871: gap of 50 bp
            *    64872    74540: contig of 9669 bp in length
            *    74541    74590: gap of 50 bp
            *    74591    90360: contig of 15770 bp in length
            *    90361    90833: gap of 473 bp
            *    90834   100046: contig of 9213 bp in length
            *   100047   100096: gap of 50 bp
            *   100097   104650: contig of 4554 bp in length
            *   104651   104950: gap of 300 bp
            *   104951   106450: contig of 1500 bp in length
            *   106451   106608: gap of 158 bp
            *   106609   109787: contig of 3179 bp in length
```

```
          *   109788   109837: gap of 50 bp
          *   109838   121211: contig of 11374 bp in length
          *   121212   121261: gap of 50 bp
          *   121262   131158: contig of 9897 bp in length
          *   131159   131208: gap of 50 bp
          *   131209   139112: contig of 7904 bp in length
          *   139113   139162: gap of 50 bp
          *   139163   145938: contig of 6776 bp in length
          *   145939   145988: gap of 50 bp
          *   145989   149971: contig of 3983 bp in length
          *   149972   150021: gap of 50 bp
          *   150022   154124: contig of 4103 bp in length
          *   154125   154174: gap of 50 bp
          *   154175   156468: contig of 2294 bp in length
          *   156469   156790: gap of 322 bp
          *   156791   171546: contig of 14756 bp in length
          *   171547   171596: gap of 50 bp
          *   171597   175250: contig of 3654 bp in length
          *   175251   175300: gap of 50 bp
          *   175301   190467: contig of 15167 bp in length
          *   190468   190517: gap of 50 bp
          *   190518   195552: contig of 5035 bp in length
          *   195553   195602: gap of 50 bp
          *   195603   198111: contig of 2509 bp in length
          *   198112   199392: gap of 1281 bp
          *   199393   202853: contig of 3461 bp in length
          *   202854   202903: gap of 50 bp
          *   202904   207061: contig of 4158 bp in length
          *   207062   207111: gap of 50 bp
          *   207112   212715: contig of 5604 bp in length
          *   212716   212765: gap of 50 bp
          *   212766   251666: contig of 38901 bp in length
          *   251667   251716: gap of 50 bp
          *   251717   258508: contig of 6792 bp in length
          *   258509   259306: gap of 798 bp
          *   259307   263777: contig of 4471 bp in length
          *   263778   263901: gap of 124 bp
          *   263902   275868: contig of 11967 bp in length
          *   275869   275968: gap of unknown length
          *   275969   277094: contig of 1126 bp in length
          *   277095   277194: gap of unknown length
          *   277195   278211: contig of 1017 bp in length
          *   278212   278311: gap of unknown length
          *   278312   279534: contig of 1223 bp in length
          *   279535   279634: gap of unknown length
          *   279635   280658: contig of 1024 bp in length
          *   280659   280758: gap of unknown length
          *   280759   282222: contig of 1464 bp in length
          *   282223   282322: gap of unknown length
          *   282323   283471: contig of 1149 bp in length
          *   283472   283571: gap of unknown length
          *   283572   284998: contig of 1427 bp in length
          *   284999   285098: gap of unknown length
          *   285099   286701: contig of 1603 bp in length
          *   286702   286801: gap of unknown length
          *   286802   288578: contig of 1777 bp in length
          *   288579   288678: gap of unknown length
          *   288679   290266: contig of 1588 bp in length

  Query Match             63.0%;  Score 25.2;  DB 14;  Length 303308;
  Best Local Similarity   90.0%;  Pred. No. 16;
  Matches   27;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 GATACCTACTCCCAAACAATCCATCTTACT 32
              | |||||||||||||||||| ||| |||||
Db     172671 GCTACCTACTCCCAAACAATACATATTACT 172700


RESULT 13
AC084290/c
LOCUS       AC084290               55001 bp    DNA     linear   PRI 28-MAR-2002
DEFINITION  Homo sapiens 12 BAC RP11-568G5 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC084290
VERSION     AC084290.15  GI:19774277
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 55001)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
            Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
            Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
            Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P.,
            Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M.,
            Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S.,
            Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
            Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
            King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
            Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O.,
            Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X.,
            Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P.,
            Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
            Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S.,
            Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
            Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
            Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 55001)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 55001)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   4  (bases 1 to 55001)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Mar 28, 2002 this sequence version replaced gi:19747063.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
```

gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT.

```
FEATURES             Location/Qualifiers
     source          1. .55001
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="RP11-568G5"
     misc_feature    1. .1301
                     /note="overlaps bases 177088. .178388 of clone AC079630"
                     /function="clone overlap"
     repeat_region   35. .229
                     /rpt_family="MIR"
     repeat_region   1002. .1023
                     /rpt_family="AT_rich"
     repeat_region   1092. .1227
                     /rpt_family="MIR"
     repeat_region   complement(2156. .2209)
                     /rpt_family="L2"
     repeat_region   3563. .3650
                     /rpt_family="(CCGTA)n"
     repeat_region   complement(3688. .3765)
                     /rpt_family="L2"
     repeat_region   complement(3840. .3908)
                     /rpt_family="L1MD3"
     repeat_region   complement(4234. .4464)
                     /rpt_family="MIR"
     repeat_region   complement(4847. .4914)
                     /rpt_family="MIR"
     repeat_region   4938. .5053
                     /rpt_family="L2"
     repeat_region   6350. .6660
                     /rpt_family="AluSg"
     repeat_region   complement(7309. .7743)
                     /rpt_family="LTR43"
     repeat_region   7744. .9144
                     /rpt_family="L1PA4"
     repeat_region   complement(9151. .9317)
                     /rpt_family="LTR43"
     repeat_region   10587. .10846
                     /rpt_family="L1M4"
     repeat_region   10912. .11015
                     /rpt_family="L1P"
     repeat_region   11098. .11271
                     /rpt_family="L1ME"
     repeat_region   11450. .11635
                     /rpt_family="MLT1I"
     repeat_region   11941. .12179
                     /rpt_family="MIR"
     repeat_region   12797. .12828
                     /rpt_family="(CATATA)n"
     repeat_region   complement(12975. .13087)
                     /rpt_family="FLAM_C"
     repeat_region   13714. .13737
                     /rpt_family="AT_rich"
     repeat_region   14058. .14219
                     /rpt_family="MER45"
     repeat_region   14411. .14545
                     /rpt_family="MIR"
     repeat_region   complement(15058. .15127)
                     /rpt_family="MIR"
     repeat_region   16186. .16413
                     /rpt_family="MIR"
     repeat_region   complement(17624. .17922)
                     /rpt_family="AluJb"
     repeat_region   complement(18309. .18606)
                     /rpt_family="AluSx"
     repeat_region   18844. .18854
                     /rpt_family="AT_rich"
     repeat_region   complement(18855. .19135)
                     /rpt_family="AluJb"
     repeat_region   19136. .19158
                     /rpt_family="AT_rich"
     repeat_region   complement(20660. .20753)
                     /rpt_family="MER91C"
     repeat_region   22179. .22680
                     /rpt_family="(TAAAA)n"
     repeat_region   22180. .22358
                     /rpt_family="AluJo"
     repeat_region   22359. .22632
                     /rpt_family="AluJb"
     repeat_region   22633. .22680
                     /rpt_family="(TAAAA)n"
     repeat_region   22682. .22789
                     /rpt_family="AluJo"
     repeat_region   22796. .23099
                     /rpt_family="AluY"
     repeat_region   complement(23147. .23327)
```

```
  Query Match           60.0%;  Score 24;  DB 8;  Length 55001;
  Best Local Similarity   84.4%;  Pred. No. 62;
  Matches   27;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCAT 35
              | || |||||||||||||||| || | ||||||
Db      47598 AAACTTACTCCCAAACAATCTATGTCACTCAT 47567


RESULT 14
AL807784/c
LOCUS       AL807784              127196 bp    DNA     linear   ROD 13-NOV-2002
DEFINITION  Mouse DNA sequence from clone RP23-448C18 on chromosome X, complete
            sequence.
ACCESSION   AL807784
VERSION     AL807784.11  GI:25045332
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
```

```
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 127196)
  AUTHORS   Howden,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 15, 2002 this sequence version replaced gi:24939941.
            Sequence from the Mouse Genome Sequencing Consortium whole genome
            shotgun may have been used to confirm this sequence.  Sequence data
            from the whole genome shotgun alone has only been used where it has
            a phred quality of at least 30.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-448C18 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES              Location/Qualifiers
     source           1. .127196
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:10090"
                      /chromosome="X"
                      /clone="RP23-448C18"
                      /clone_lib="RPCI-23"
ORIGIN

  Query Match             60.0%;  Score 24;  DB 9;  Length 127196;
  Best Local Similarity   75.0%;  Pred. No. 55;
  Matches   30;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||| ||  || || ||| |  ||||||| |||| ||
Db     110413 AGGATACATAAGCCAAATCAACCTTTCTTACTAATGCCAC 110374


RESULT 15
AC157528
LOCUS       AC157528              154857 bp    DNA     linear   PRI 27-APR-2005
DEFINITION  Pan troglodytes BAC clone CH251-422M6 from chromosome unknown,
            complete sequence.
ACCESSION   AC157528
VERSION     AC157528.2  GI:62000959
KEYWORDS    HTG.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1  (bases 1 to 154857)
  AUTHORS   Shah,N., Cotton,M. and Lewis,S.
  TITLE     The sequence of Pan troglodytes BAC clone CH251-422M6
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 154857)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-FEB-2005) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 154857)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2005) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE   4  (bases 1 to 154857)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2005) Washington University School of Medicine,
            Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO
            63108, USA
COMMENT     On Mar 30, 2005 this sequence version replaced gi:59933404.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: C_AB0422M06


            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The CHORI-251 Chimpanzee BAC library has been constructed at the
            Children's Hospital Oakland Research Institute, BACPAC Resources,
            by Dr. Baoli Zhu.  DNA was isolated from white blood cells obtained
            from a male chimpanzee (Pan troglodytes, 'Clint', Yerkes #C0471;
            birthdate:6-6-80).  The clone and detailed information can be
            obtained from Pieter de Jong and co-workers at
            http://www.bacpac.chori.org.

            This sequence is the entire insert of the clone.
FEATURES              Location/Qualifiers
     source           1. .154857
                      /organism="Pan troglodytes"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9598"
                      /chromosome="unknown"
                      /clone="CH251-422M6"
                      /clone_lib="CHORI251"
     unsure           74900
                      /note="Unresolved bases"
     unsure           74903
                      /note="Unresolved bases"
     unsure           127657. .127881
                      /note="Sequence derived from one plasmid subclone."
ORIGIN

  Query Match             60.0%;  Score 24;  DB 8;  Length 154857;
```

Best Local Similarity 84.4%; Pred. No. 53;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCAT 35
              | || ||||||||||||||| || | ||||||
Db     123632 AAACTTACTCCCAAACAATCTATGTCACTCAT 123663
```

Search completed: March 5, 2006, 21:55:48
Job time : 176.689 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:08:23 ; Search time 26.619 Seconds
(without alignments)
10014.946 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278
Perfect score: 40
Sequence: 1 aggatacctactcccaaacaatccatcttactcatgcaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	100.0	40	5	AAH76337	Aah76337 Z. mays M
2	40	100.0	50	5	AAH76336	Aah76336 Z. mays M
3	40	100.0	158	5	AAH76334	Aah76334 Z. mays M
4	40	100.0	255	5	AAH76340	Aah76340 Z. mays M
5	40	100.0	1394	2	AAX07408	Aax07408 Zea mays
6	40	100.0	1394	2	AAX07409	Aax07409 Zea mays
7	40	100.0	1394	5	AAH76332	Aah76332 Z. mays M
8	40	100.0	1394	5	AAH76333	Aah76333 Z. mays M
9	30.4	76.0	2000	11	ACL38730	Acl38730 Rice stre
10	22.4	56.0	226215	11	ACN45146	Acn45146 Human gen
c 11	21.6	54.0	4110	8	ABZ10202	Abz10202 Haematopo
c 12	21.6	54.0	6309	6	ABL32304	Abl32304 Human imm
c 13	21.6	54.0	7110	10	ADB54282	Adb54282 Pretreate
c 14	21.6	54.0	7110	10	ADE84196	Ade84196 Human lym
c 15	21.6	54.0	7110	13	ADS89580	Ads89580 Oligonucl
c 16	21.4	53.5	1341	6	ABZ13750	Abz13750 Arabidops
c 17	21.4	53.5	6681	6	ABL32155	Abl32155 Human imm
c 18	21.4	53.5	6681	6	ABL54304	Abl54304 Chemicall
c 19	21.4	53.5	8168	6	ABL32928	Abl32928 Human imm
c 20	21.4	53.5	8168	6	AAS63328	Aas63328 Chemicall
c 21	21.4	53.5	24401	4	ABL03396	Abl03396 Drosophil
22	21	52.5	1531	14	ADW17938	Adw17938 Eucalyptu
23	21	52.5	2000	11	ACL35665	Acl35665 Rice stre
24	21	52.5	2000	12	ADJ41594	Adj41594 Plant cDN
c 25	21	52.5	2792	12	ADM98941	Adm98941 Diterpene
c 26	21	52.5	3117	12	ADM98942	Adm98942 Diterpene
c 27	21	52.5	5507	8	ABZ10204	Abz10204 Haematopo
c 28	21	52.5	5507	8	ABZ10058	Abz10058 Haematopo
c 29	21	52.5	8404	4	AAS46500	Aas46500 Tumour su
c 30	21	52.5	8404	6	ABL33595	Abl33595 Human imm
c 31	21	52.5	8404	10	ADB54216	Adb54216 Pretreate
c 32	21	52.5	8404	10	ADB54088	Adb54088 Pretreate
c 33	21	52.5	8404	10	ADE84076	Ade84076 Human lym
c 34	21	52.5	8404	10	ADE84152	Ade84152 Human lym
c 35	21	52.5	8404	13	ADS89242	Ads89242 Oligonucl
c 36	21	52.5	8404	13	ADS89516	Ads89516 Oligonucl
c 37	21	52.5	10286	4	AAS45308	Aas45308 Chemicall
c 38	21	52.5	10286	6	ABK28147	Abk28147 DNA trans
39	21	52.5	63115	10	ADC85419	Adc85419 Mouse Tle
40	21	52.5	63294	9	ADA02939	Ada02939 Mouse Tle
41	21	52.5	63294	10	ADB72677	Adb72677 Mouse Tle
42	21	52.5	63294	12	ADM74534	Adm74534 Murine ca
c 43	21	52.5	106315	11	ACN43966	Acn43966 Human gen
c 44	20.8	52.0	891	8	ABZ51893	Abz51893 Aspergill
c 45	20.8	52.0	6620	4	AAS45488	Aas45488 Chemicall

ALIGNMENTS

RESULT 1
AAH76337
ID AAH76337 standard; DNA; 40 BP.
XX
AC AAH76337;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 14; Page 32; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The

```
CC   present sequence represents a DNA fragment upstream of the TATA box of a
CC   Z. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX
SQ   Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 40;  DB 5;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40


RESULT 2
AAH76336
ID   AAH76336 standard; DNA; 50 BP.
XX
AC   AAH76336;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 5; Page 47; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a DNA fragment -72 to -111 bases upstream of
CC   the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC   nucleotide sequence
XX
SQ   Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 40;  DB 5;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db         11 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 50


RESULT 3
AAH76334
ID   AAH76334 standard; DNA; 158 BP.
XX
AC   AAH76334;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 5; Page 47; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a DNA fragment -38 to -195 bases upstream of
CC   the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC   nucleotide sequence
XX
SQ   Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 40;  DB 5;  Length 158;
  Best Local Similarity   100.0%;  Pred. No. 5e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db         86 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 125


RESULT 4
AAH76340
ID   AAH76340 standard; DNA; 255 BP.
XX
AC   AAH76340;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 promoter fragment.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; promoter; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
```

```
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Example 5; Fig 8; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a Z. mays Ms45 promoter fragment
XX
SQ   Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 40;  DB 5;  Length 255;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db         99 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 138


RESULT 5
AAX07408
ID   AAX07408 standard; DNA; 1394 BP.
XX
AC   AAX07408;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW   Ms45; male; tissue-preferred; regulatory region; plant cells;
KW   plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS   Zea mays.
XX
PN   WO9859061-A1.
XX
PD   30-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US012895.
XX
PR   23-JUN-1997;   97US-00880499.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman GA,  Kendall TL;
XX
DR   WPI; 1999-105628/09.
XX
PT   New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT   - useful in mediating plant fertility, especially hybrid seed production.
XX
PS   Claim 2; Page 22-23; 39pp; English.
XX
CC   The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC   region. It may be used in the construction of a vector for a method of
CC   producing exogenous genes in a male tissue- preferred manner, which is
CC   useful in restoring or conferring fertility, such as in hybrid seed
CC   production. In conferring fertility, a monocot/dicot plant is transformed
CC   with the exogenous nucleotide sequence (a male sterility gene, preferably
CC   Ms45), which encodes a product selected from auxins, rolB and diptheria
CC   toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 40;  DB 2;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 6
AAX07409
ID   AAX07409 standard; DNA; 1394 BP.
XX
AC   AAX07409;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW   Ms45; male; tissue-preferred; regulatory region; plant cells;
KW   plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS   Zea mays.
XX
PN   WO9859061-A1.
XX
PD   30-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US012895.
XX
PR   23-JUN-1997;   97US-00880499.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman GA,  Kendall TL;
XX
DR   WPI; 1999-105628/09.
XX
PT   New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT   - useful in mediating plant fertility, especially hybrid seed production.
XX
PS   Claim 3; Page 23-24; 39pp; English.
XX
CC   The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC   region. It may be used in the construction of a vector for a method of
CC   producing exogenous genes in a male tissue- preferred manner, which is
CC   useful in restoring or conferring fertility, such as in hybrid seed
CC   production. In conferring fertility, a monocot/dicot plant is transformed
CC   with the exogenous nucleotide sequence (a male sterility gene, preferably
CC   Ms45), which encodes a product selected from auxins, rolB and diptheria
CC   toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 40;  DB 2;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278
```

```
RESULT 7
AAH76332
ID   AAH76332 standard; DNA; 1394 BP.
XX
AC   AAH76332;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 4; Page 46; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a nucleic acid sequence encoding an Ms45 male
CC   -tissue preferred regulatory region from Z. mays
XX
SQ   Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 40;  DB 5;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 8
AAH76333
ID   AAH76333 standard; DNA; 1394 BP.
XX
AC   AAH76333;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 4; Page 47; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a nucleic acid sequence encoding an Ms45 male
CC   -tissue preferred regulatory region from Z. mays
XX
SQ   Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 40;  DB 5;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 9
ACL38730
ID   ACL38730 standard; cDNA; 2000 BP.
XX
AC   ACL38730;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Rice stress-regulated promoter SEQ ID NO:17293.
XX
KW   ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW   agriculture.
XX
OS   Oryza sativa.
XX
PN   WO2003008540-A2.
XX
PD   30-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019668.
XX
PR   22-JUN-2001; 2001US-0300112P.
PR   24-AUG-2001; 2001US-0314662P.
PR   26-SEP-2001; 2001US-0325277P.
PR   21-NOV-2001; 2001US-0332132P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Kreps J,  Briggs SP,  Cooper B,  Glazebrook J,  Goff SA,  Katagiri F;
PI   Moughamer T,  Provart N,  Ricke D,  Zhu T;
XX
DR   WPI; 2003-248011/24.
XX
```

PT New stress-responsive nucleic acid, useful for altering the
PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT stress, salt stress or osmotic stress.
XX
PS Claim 48; SEQ ID NO 17293; 89pp; English.
XX
CC The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention
XX
SQ Sequence 2000 BP; 616 A; 370 C; 355 G; 659 T; 0 U; 0 Other;

Query Match 76.0%; Score 30.4; DB 11; Length 2000;
Best Local Similarity 85.0%; Pred. No. 0.047;
Matches 34; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
|| | |||||||||||||||||| || ||||||||||
Db 1840 AGAACACCTACTCCCAAACAATCACGCTGACTCATGCAAC 1879

RESULT 10
ACN45146
ID ACN45146 standard; DNA; 226215 BP.
XX
AC ACN45146;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG1639824.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1948; 0pp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 226215 BP; 65331 A; 42804 C; 45218 G; 72862 T; 0 U; 0 Other;

Query Match 56.0%; Score 22.4; DB 11; Length 226215;
Best Local Similarity 72.5%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
|| || ||| || ||||||| |||||||||| | ||
Db 183206 AGAATTCCTGCTTCCAAACACAGCATCTTACTCCTTTTAC 183245

RESULT 11
ABZ10202/c
ID ABZ10202 standard; DNA; 4110 BP.
XX
AC ABZ10202;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #342.
XX
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwope I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
PS Claim 28; SEQ ID NO 342; 117pp; English.
XX
CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA

```
CC   sequences. The nucleotide sequences from the present invention can also
CC   be used for detecting a predisposition to, differentiation between
CC   subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC   haematopoietic cell proliferative disorders. The present method enables a
CC   highly specific classification of haematopoietic cell proliferative
CC   disorders allowing for improved and informed treatment of patients
XX
SQ   Sequence 4110 BP; 945 A; 0 C; 1132 G; 2033 T; 0 U; 0 Other;

  Query Match             54.0%;  Score 21.6;  DB 8;  Length 4110;
  Best Local Similarity   85.7%;  Pred. No. 1.6e+02;
  Matches   24;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTAC 31
              | |||||  ||||||||||||||| |||
Db       3083 AAACCTAACCCCAAACAATCCATCCTAC 3056


RESULT 12
ABL32304/c
ID   ABL32304 standard; DNA; 6309 BP.
XX
AC   ABL32304;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 277.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 277; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 6309 BP; 1758 A; 69 C; 1364 G; 3118 T; 0 U; 0 Other;

  Query Match             54.0%;  Score 21.6;  DB 6;  Length 6309;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+02;
  Matches   27;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              ||||||| ||||| || | | |||| | | || |||
Db       3335 ATACCTATTCCCATACGACCAATCTCAATAATCCAA 3300


RESULT 13
ADB54282/c
ID   ADB54282 standard; DNA; 7110 BP.
XX
AC   ADB54282;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Pretreated genomic DNA region 206.
XX
KW   colon cell proliferative disorder; non methylated CpG dinucleotide;
KW   cytostatic; cancer; adenoma; carcinoma; cytosine methylation state; ds.
XX
OS   Unidentified.
XX
PN   WO2003072821-A2.
XX
PD   04-SEP-2003.
XX
PF   27-FEB-2003; 2003WO-EP002035.
XX
PR   27-FEB-2002; 2002EP-00004551.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Adorjan P,  Burger M,  Maier S,  Nimmrich I,  Becker E,  Lesche R;
PI   Rujan T,  Schmitt A;
XX
DR   WPI; 2003-731620/69.
XX
PT   Detecting and differentiating between colon cell proliferative disorders
PT   associated with a gene or its regulatory regions comprises contacting a
PT   target nucleic acid in a biological sample obtained from the subject with
PT   a reagent.
XX
PS   Claim 32; SEQ ID NO 338; 74pp; English.
XX
CC   The invention relates to a novel method for detecting and differentiating
CC   between colon cell proliferative disorders associated with at least one
CC   gene or its regulatory regions. The method comprises contacting a target
CC   nucleic acid in a biological sample obtained from the subject with at
CC   least one reagent or a series of reagents, where the reagent or series of
CC   reagents, distinguishes between methylated and non methylated CpG
CC   dinucleotides within the target nucleic acid. The molecules of the
CC   invention demonstrate cytostatic activity whilst the method may useful
CC   for detecting and differentiating between colon cell proliferative
CC   disorders, including cancers such as colon adenoma and colon carcinoma.
CC   The PNA (peptide nucleic acid)-oligomers are useful as probes for
CC   determining cytosine methylation state or single nucleotide
CC   polymorphisms. The current sequence is that of the pretreated genomic DNA
CC   region of the invention. This sequence is not shown within the
CC   specification but is taken from Wipoweb.
XX
SQ   Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

  Query Match             54.0%;  Score 21.6;  DB 10;  Length 7110;
  Best Local Similarity   85.7%;  Pred. No. 1.7e+02;
  Matches   24;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTAC 31
              | |||||  ||||||||||||||| |||
Db       3083 AAACCTAACCCCAAACAATCCATCCTAC 3056


RESULT 14
ADE84196/c
```

ID ADE84196 standard; DNA; 7110 BP.
XX
AC ADE84196;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human lymphoid cell proliferative disorder gene derived DNA #132.
XX
KW ds; lymphoid cell proliferative disorder; methylation;
KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
KW diffuse large B-cell lymphoma; mantle cell lymphoma;
KW chronic lymphocytic leukemia; small lymphocytic lymphoma;
KW follicular lymphoma; diagnosis; prognosis.
XX
OS Homo sapiens.
XX
PN WO2003044226-A2.
XX
PD 30-MAY-2003.
XX
PF 25-NOV-2002; 2002WO-EP013265.
XX
PR 23-NOV-2001; 2001DE-01057491.
PR 28-DEC-2001; 2001DE-01064501.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimmrich I;
XX
DR WPI; 2003-457621/43.
XX
PT Detecting and differentiating between lymphoid cell proliferative
PT disorders comprises contacting a target nucleic acid with at least one
PT reagent that distinguishes between methylated and non-methylated CpG
PT dinucleotides.
XX
PS Claim 26; SEQ ID NO 192; 448pp; English.
XX
CC The invention relates to a method of detecting and differentiating
CC between lymphoid cell proliferative disorders associated with at least
CC one gene and/or their regulatory regions in a subject by contacting a
CC target nucleic acid in a biological sample obtained from the subject with
CC at least one reagent or series of reagents that distinguish between
CC methylated and non-methylated CpG dinucleotides within the target nucleic
CC acid. The genes and/or their regulatory regions are preferably selected
CC from MDR1, CSNK2B, EGR4, AR, CDK4, RB2, CDC25A, GPIb beta, MYOD1, CDH3,
CC MYCL1, ELK1, ABL1, APC, BCL2, CDH1, CDKN1A, CDKN1B, CDKN2a, CDKN2B, FOS,
CC GSTP1, HIC-1, MGMT, MLH1, MOS, MYC, PTEN, RBL2, TGFBR2, TP73, CDKNIC,
CC GSK3beta, ESR1, APAF1, BAK1, BAX or HOXA5. Oligomers, peptide nucleic
CC acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences
CC of the genes are useful for detecting the methylation state of all the
CC CpG dinucleotides within one or more the sequences, or their complements,
CC for determining the cytosine methylation state and or single nucleotide
CC polymorphisms (SNPs), and for differentiating at least two of the medical
CC conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma,
CC chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular
CC lymphoma. They are also useful for detecting of a predisposition to,
CC differentiation between subclasses, diagnosis, prognosis, treating and/or
CC monitoring of lymphoid cell proliferative disorder. This sequence
CC represents a nucleic acid of a pretreated genomic DNA derived from the
CC above mentioned genes.
XX
SQ Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 10; Length 7110;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 ATACCTACTCCCAAACAATCCATCTTAC 31
| ||||| ||||||||||||||| |||
Db 3083 AAACCTAACCCCAAACAATCCATCCTAC 3056

RESULT 15
ADS89580/c
ID ADS89580 standard; DNA; 7110 BP.
XX
AC ADS89580;
XX
DT 18-NOV-2004 (first entry)
XX
DE Oligonucleotide of the invention SEQ ID NO:596.
XX
KW ss; cell proliferative disorder; breast; methylation; cytostatic;
KW gene therapy; single nucleotide polymorphism; SNP.
XX
OS Unidentified.
XX
PN WO2004035803-A2.
XX
PD 29-APR-2004.
XX
PF 01-OCT-2003; 2003WO-EP010881.
XX
PR 01-OCT-2002; 2002DE-01045779.
PR 07-JAN-2003; 2003DE-01000096.
PR 17-APR-2003; 2003DE-01017955.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Foekens J, Harbeck N, Koenig T, Maier S, Martens J, Model F;
PI Nimmrich I, Rujan T, Schmitt A, Schmitt M, Look MP, Marx A;
XX
DR WPI; 2004-348468/32.
XX
PT Predicting responsiveness of a subject with breast cell proliferative
PT disorder, useful for treating or differentiating breast cell
PT proliferative disorders comprises analyzing methylation pattern of a
PT genomic DNA from the subject.
XX
PS Disclosure; SEQ ID NO 596; 104pp; English.
XX
CC The invention relates to a novel method for predicting the responsiveness
CC of a subject with a cell proliferative disorder of the breast tissues to
CC a therapy comprising analysing the methylation pattern of a target
CC nucleic acid by contacting at least one of the target nucleic acids in a
CC biological sample obtained from the subject prior to or during treatment.
CC The method of the invention has cytostatic activity, and may have a use
CC in gene therapy. The set of oligonucleotides comprising at least two of
CC the oligomers are useful for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The
CC methods, nucleic acid, oligonucleotide, and kit are useful for the
CC treatment, characterisation, classification and/or differentiation, of
CC breast cell proliferative disorders. The method is also useful for
CC predicting the responsiveness of a subject with a cell proliferative
CC disorder of the breast tissues to a therapy. The present sequence is used
CC in the exemplification of the invention.
XX
SQ Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 13; Length 7110;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 ATACCTACTCCCAAACAATCCATCTTAC 31
| ||||| ||||||||||||||| |||
Db 3083 AAACCTAACCCCAAACAATCCATCCTAC 3056

Search completed: March 5, 2006, 18:11:42
Job time : 26.619 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 18:11:59 ; Search time 188.843 Seconds
(without alignments)
9910.279 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278
Perfect score: 40
Sequence: 1 aggatacctactcccaaacaatccatcttactcatgcaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	40	100.0	687	9	CC656939	CC656939 OGWDQ20TV
	2	40	100.0	915	10	CG224225	CG224225 OG1AG08TV
	3	40	100.0	963	9	CC656933	CC656933 OGWDQ20TM
	4	35.2	88.0	702	10	CW324514	CW324514 104_819_1
	5	24.2	60.5	296	10	CW445575	CW445575 fsbb001f1
	6	23.8	59.5	528	8	H24137	H24137 ym50g08.r1
c	7	23.6	59.0	375	9	BZ706605	BZ706605 SM416B1-G
	8	23.2	58.0	738	7	CO117589	CO117589 GR__Eb01H
	9	23.2	58.0	783	11	CR181951	CR181951 Reverse s
c	10	23.2	58.0	1018	10	CL109241	CL109241 ISB1-51D1
	11	23.2	58.0	1253	6	CD495749	CD495749 CDA18-D05
c	12	23	57.5	405	9	AZ883408	AZ883408 RPCI-23-1
	13	23	57.5	662	7	CJ037346	CJ037346
	14	23	57.5	790	7	CJ035837	CJ035837 CJ035837
c	15	23	57.5	798	2	BG506830	BG506830 601861443
	16	23	57.5	968	2	BF143726	BF143726 601789606
	17	22.8	57.0	771	3	BM406466	BM406466 EST580781
c	18	22.8	57.0	1044	8	DN570622	DN570622 92287550
	19	22.6	56.5	398	10	CW888409	CW888409 RPCI42_10
	20	22.6	56.5	410	1	AW257178	AW257178 EST305315
	21	22.6	56.5	560	2	BE205271	BE205271 EST397947
c	22	22.6	56.5	624	11	CR843833	CR843833 GR0AAA78D
c	23	22.6	56.5	630	9	BZ173201	BZ173201 CH230-369
c	24	22.6	56.5	644	11	CR843905	CR843905 GR0AAA79A
	25	22.6	56.5	664	9	AZ400686	AZ400686 1M0167P01
	26	22.6	56.5	691	1	AW687128	AW687128 NF006C08R
c	27	22.6	56.5	736	11	CR818130	CR818130 GR0AAA43D
	28	22.6	56.5	1080	10	CL058859	CL058859 CH216-88P
	29	22.4	56.0	298	2	BB264116	BB264116 BB264116
c	30	22.4	56.0	804	8	DT054853	DT054853 AGENCOURT
c	31	22.4	56.0	940	7	CK864514	CK864514 AGENCOURT
c	32	22.4	56.0	940	7	CK871857	CK871857 AGENCOURT
c	33	22.4	56.0	951	7	CN317130	CN317130 AGENCOURT
	34	22.2	55.5	273	1	AI206216	AI206216 qr27f06.x
	35	22.2	55.5	346	1	AI698201	AI698201 wa67g11.x
	36	22.2	55.5	398	1	AA725064	AA725064 ai06h07.s
	37	22.2	55.5	417	1	AI467879	AI467879 tj78e09.x
	38	22.2	55.5	479	2	BF601210	BF601210 266128 MA
	39	22.2	55.5	491	1	AI382034	AI382034 te33g05.x
c	40	22.2	55.5	609	3	BM027586	BM027586 GIT000109
	41	22.2	55.5	727	7	CK949745	CK949745 4074916 B
c	42	22.2	55.5	802	10	CG811817	CG811817 FSAAM35TR
	43	22.2	55.5	874	10	CZ727957	CZ727957 OC__Ba005
c	44	22.2	55.5	940	10	CG383910	CG383910 OGZAP61TV
	45	22.2	55.5	1090	10	CL084441	CL084441 ISB1-3G16

ALIGNMENTS

RESULT 1
CC656939/c
LOCUS CC656939 687 bp DNA linear GSS 19-JUN-2003
DEFINITION OGWDQ20TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15, genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 687)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGWDQ20TM
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: methylation filtered.
FEATURES Location/Qualifiers
source 1. .687
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0554D15"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 100.0%; Score 40; DB 9; Length 687;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40

```
                     ||||||||||||||||||||||||||||||||||||||||
Db        337 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 298


RESULT 2
CG224225
LOCUS       CG224225                915 bp    DNA     linear   GSS 22-AUG-2003
DEFINITION  OG1AG08TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0716B15,
            genomic survey sequence.
ACCESSION   CG224225
VERSION     CG224225.1  GI:34124113
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1AG08TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0716B15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             100.0%;  Score 40;  DB 10;  Length 915;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db        618 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 657


RESULT 3
CC656933
LOCUS       CC656933                963 bp    DNA     linear   GSS 19-JUN-2003
DEFINITION  OGWDQ20TM ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15,
            genomic survey sequence.
ACCESSION   CC656933
VERSION     CC656933.1  GI:32060225
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 963)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OGWDQ20TV
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .963
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0554D15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             100.0%;  Score 40;  DB 9;  Length 963;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-05;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db        836 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 875


RESULT 4
CW324514
LOCUS       CW324514                702 bp    DNA     linear   GSS 31-OCT-2004
DEFINITION  104_819_11477203_148_35910_078 Sorghum methylation filtered library
            (LibID: 104) Sorghum bicolor genomic clone 11477203, genomic survey
            sequence.
ACCESSION   CW324514
VERSION     CW324514.1  GI:55040702
KEYWORDS    GSS.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
            Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
            McMenamy,J., Smith,M., Holeman,H., Roe,B.A, Wiley,G., Korf,I.F.,
            Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
            Martienssen,R.A.
  TITLE     Sorghum genome sequencing by methylation filtration
  JOURNAL   PLoS Biol. 3 (1), e13 (2005)
   PUBMED   15660154
COMMENT     Contact: Bedell JA
            Orion Genomics, LLC
            4041 Forest Park Ave, St. Louis, MO 63108, USA
            Tel: 314 615 6979
            Fax: 314 615 5975
            Email: jbedell@oriongenomics.com
            Plate: 819  row: c  column: 19
            Seq primer: SWfor Forward
            Class: methylation filtered
            High quality sequence stop: 702.
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="Sorghum bicolor"
                     /mol_type="genomic DNA"
                     /cultivar="ATx623"
                     /db_xref="taxon:4558"
                     /clone="11477203"
                     /clone_lib="Sorghum methylation filtered library (LibID:
                     104)"
                     /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
                     prepared from purified nuclei was randomly sheared,
```

end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."

ORIGIN

Query Match 88.0%; Score 35.2; DB 10; Length 702;
Best Local Similarity 92.5%; Pred. No. 0.0034;
Matches 37; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              |||| |||||||||||||||||||| |||| ||||||||
Db         74 AGGACACCTACTCCCAAACAATCCATGTTACCCATGCAAC 113
```

RESULT 5
CW445575
LOCUS CW445575 296 bp DNA linear GSS 02-NOV-2004
DEFINITION fsbb001f170m16k0 Sorghum methylation filtered library (LibID: 104) Sorghum bicolor genomic clone fsbb001f170m16, genomic survey sequence.
ACCESSION CW445575
VERSION CW445575.1 GI:55193536
KEYWORDS GSS.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 296)
AUTHORS Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D., Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K., McMenamy,J., Smith,M., Holeman,H., Roe,B.A, Wiley,G., Korf,I.F., Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and Martienssen,R.A.
TITLE Sorghum genome sequencing by methylation filtration
JOURNAL PLoS Biol. 3 (1), e13 (2005)
PUBMED 15660154
COMMENT Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fsbb001f170 row: m column: 16
Seq primer: k Reverse
Class: methylation filtered
High quality sequence stop: 296.

FEATURES Location/Qualifiers
source 1. .296
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clone="fsbb001f170m16"
/clone_lib="Sorghum methylation filtered library (LibID: 104)"
/note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."

ORIGIN

Query Match 60.5%; Score 24.2; DB 10; Length 296;
Best Local Similarity 90.2%; Pred. No. 75;
Matches 37; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
Qy          1 AGGATACCTACTCCCAAACAAT-CCATCTTACTCATGCAAC 40
              |||| |||||||||||||||| |||| |||| ||||||||
Db        242 AGGACACCTACTCCCAAACAATCCCATGTTACCCATGCAAC 282
```

RESULT 6
H24137
LOCUS H24137 528 bp mRNA linear EST 06-JUL-1995
DEFINITION ym50g08.r1 Soares infant brain 1NIB Homo sapiens cDNA clone IMAGE:51939 5', mRNA sequence.
ACCESSION H24137
VERSION H24137.1 GI:892832
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 528)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1161
High quality sequence stops: 349
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1161 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 349.

FEATURES Location/Qualifiers
source 1. .528
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:424875"
/db_xref="taxon:9606"
/clone="IMAGE:51939"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain 1NIB"
/note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not I; Site_2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 59.5%; Score 23.8; DB 8; Length 528;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCA 38
              | | |||||| |||||| | ||||||| ||||||
Db        346 AAATCTACTCTCAAACATTTCATCTTATACATGCA 380
```

RESULT 7
BZ706605/c
LOCUS BZ706605 375 bp DNA linear GSS 18-MAY-2003
DEFINITION SM416B1-G07_55_13.ab1 Spider Monkey genomic BAC library Ateles geoffroyi genomic, genomic survey sequence.
ACCESSION BZ706605

```
VERSION     BZ706605.1  GI:30843142
KEYWORDS    GSS.
SOURCE      Ateles geoffroyi (black-handed spider monkey)
  ORGANISM  Ateles geoffroyi
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
            Cebidae; Atelinae; Ateles.
REFERENCE   1  (bases 1 to 375)
  AUTHORS   Qian,Y., Jin,L. and Su,B.
  TITLE     Construction and characterization of bacterial artificial
            chromosome library of black-handed spider monkey (Ateles geoffroyi)
  JOURNAL   Genome 47 (2), 239-245 (2004)
   PUBMED   15060576
COMMENT     Contact: Qian Y, Jin L, Su B.
            Center for Genome Information
            University of Cincinnati
            Kettering Lab, 3223 Eden Ave., Cincinnati, OH 45267-0056, USA
            Tel: 1-513-558-6678
            Fax: 1-513-558-4505
            Email: subn@ucmail.uc.edu
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .375
                     /organism="Ateles geoffroyi"
                     /mol_type="genomic DNA"
                     /strain="black-handed spider monkey"
                     /db_xref="taxon:9509"
                     /sex="Male"
                     /tissue_type="Skin"
                     /cell_type="Fibroblast"
                     /cell_line="AG05352"
                     /dev_stage="3 DA"
                     /clone_lib="Spider Monkey genomic BAC library"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Genomic DNA was
                     partially digested with EcoRI, Vector, pBACe3.6,
                     Recombinants were transformed into DH10B. "
ORIGIN

  Query Match             59.0%;  Score 23.6;  DB 9;  Length 375;
  Best Local Similarity   74.4%;  Pred. No. 1.4e+02;
  Matches   29;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 GGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||| |||||    ||| || ||| ||||| | |||||
Db        255 GGATATCTACTTGTGAACCATACATTTTACTNAAGCAAC 217


RESULT 8
CO117589
LOCUS       CO117589                738 bp    mRNA    linear   EST 16-JUN-2004
DEFINITION  GR__Eb01H15.r GR__Eb Gossypium raimondii cDNA clone GR__Eb01H15 3',
            mRNA sequence.
ACCESSION   CO117589
VERSION     CO117589.1  GI:48816276
KEYWORDS    EST.
SOURCE      Gossypium raimondii
  ORGANISM  Gossypium raimondii
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE   1  (bases 1 to 738)
  AUTHORS   Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
            Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
            Wing,R.A.
  TITLE     Global assembly of Cotton ESTs
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            The University of Arizona
            Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            Plate: 01  row: H  column: 15.
FEATURES             Location/Qualifiers
     source          1. .738
                     /organism="Gossypium raimondii"
                     /mol_type="mRNA"
                     /db_xref="taxon:29730"
                     /clone="GR__Eb01H15"
                     /tissue_type="floral"
                     /dev_stage="-3 to +3 DPA"
                     /lab_host="DH10B"
                     /clone_lib="GR__Eb"
                     /note="Vector: pCMV.SPORT-6.1; Site_1: NotI; Site_2:
                     EcoRV; Library made by Invitrogen with RNA supplied by
                     Wendle lab. Directional cloned into NotI-EV. Colonies
                     plated/picked by AGI. More glycerol clones held in -80."
ORIGIN

  Query Match             58.0%;  Score 23.2;  DB 7;  Length 738;
  Best Local Similarity   77.8%;  Pred. No. 2.2e+02;
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              ||||  ||||||| ||||  ||  | ||||||||||
Db        224 ATACACACTCCCATACAACGCAAGTGACTCATGCAA 259


RESULT 9
CR181951
LOCUS       CR181951                783 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Reverse strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN361p21, genomic survey sequence.
ACCESSION   CR181951
VERSION     CR181951.1  GI:49960800
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 783)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .783
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN361p21"
                     /clone_lib="MHPN"
ORIGIN

  Query Match             58.0%;  Score 23.2;  DB 11;  Length 783;
  Best Local Similarity   77.8%;  Pred. No. 2.2e+02;
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              ||||||||||  |  ||||  ||||||||||  | ||
Db        514 ATACCTACTCTAAGCCAATGCATCTTACTCCAGGAA 549


RESULT 10
CL109241/c
LOCUS       CL109241                1018 bp    DNA     linear   GSS 05-JAN-2004
DEFINITION  ISB1-51D14_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-51D14,
            genomic survey sequence.
ACCESSION   CL109241
VERSION     CL109241.1  GI:40602876
KEYWORDS    GSS.
```

```
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1018)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000  Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 72
            High quality sequence stop: 524.
FEATURES             Location/Qualifiers
     source          1. .1018
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8364"
                     /clone="ISB1-51D14"
                     /clone_lib="ISB1"
                     /note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                     Library Segment 1"
ORIGIN

  Query Match           58.0%;  Score 23.2;  DB 10;  Length 1018;
  Best Local Similarity   77.8%;  Pred. No. 2.3e+02;
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          5 TACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||| |||||||| |||||||||| ||| | |   |||
Db        156 TACATACTCCCATACAATCCATTTTAATGAAATAAC 121


RESULT 11
CD495749
LOCUS       CD495749               1253 bp    mRNA    linear   EST 12-JUN-2003
DEFINITION  CDA18-D05.y1d-s SHGC-CDA Gasterosteus aculeatus cDNA clone
            CDA18-D05 3', mRNA sequence.
ACCESSION   CD495749
VERSION     CD495749.1  GI:31422780
KEYWORDS    EST.
SOURCE      Gasterosteus aculeatus (three spined stickleback)
  ORGANISM  Gasterosteus aculeatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
            Gasterosteidae; Gasterosteus.
REFERENCE   1  (bases 1 to 1253)
  AUTHORS   Kingsley,D.M., Peichel,C., Balabahdra,S., Grimwood,J., Dickson,M.,
            Schmutz,J. and Myers,R.M.
  TITLE     Expressed sequence tags from Gasterosteus aculeatus
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Kingsley, DM
            HHMI and Department of Developmental Biology
            Stanford University School of Medicine
            Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
            Tel: 650 725 5954
            Fax: 650 725 7739
            Email: kingsley@cmgm.stanford.edu
            Plate: 18
            High quality sequence start: 15
            High quality sequence stop: 667.
FEATURES             Location/Qualifiers
     source          1. .1253
                     /organism="Gasterosteus aculeatus"
                     /mol_type="mRNA"
                     /strain="Salinas river, CA"
                     /db_xref="taxon:69293"
                     /clone="CDA18-D05"
                     /sex="mixed male and female"
                     /tissue_type="heads and internal organs combined"
                     /dev_stage="adult"
                     /clone_lib="SHGC-CDA"
                     /note="Vector: lambda ZAP Express/pBK-CMV; Site_1: EcoR1
                     (5' adaptor); Site_2: XhoI (3' linker primer); The mixed
                     organ cDNA library was generated using the ZAP-cDNA method
                     by Stratagene. First strand cDNA synthesis was primed with
                     a a 50 bp linker primer containing an oligo dT sequence
                     preceeded by a synthetic XhoI site. 5 prime adaptors were
                     used containing an EcoR1 cohesive end. The finished cDNAs
                     were inserted in to the ZAP express vector
                     unidirectionally in the sense orientation with respect to
                     the lacZ promoter of pBK-CMV. An amplified library was
                     prepared from approximately 3 million primary clones in
                     the lambda ZAP Express vector. In vivo excision was then
                     used to generate individual pBK-CMV phagemid clones for
                     EST sequencing."
ORIGIN

  Query Match           58.0%;  Score 23.2;  DB 6;  Length 1253;
  Best Local Similarity   77.8%;  Pred. No. 2.4e+02;
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          5 TACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||| | || |||||||||||||   ||| ||||
Db       1047 TACCTCTTTCCCAACAATCCATCTTCACCATCCAAC 1082


RESULT 12
AZ883408/c
LOCUS       AZ883408               405 bp    DNA     linear   GSS 05-MAR-2001
DEFINITION  RPCI-23-189L2.TJ RPCI-23 Mus musculus genomic clone RPCI-23-189L2,
            genomic survey sequence.
ACCESSION   AZ883408
VERSION     AZ883408.1  GI:13202353
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 405)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-189L2.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 189  row: L  column: 2
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-189L2"
```

```
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
ORIGIN

  Query Match             57.5%;  Score 23;  DB 9;  Length 405;
  Best Local Similarity   74.4%;  Pred. No. 2.4e+02;
  Matches   29;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||  ||  | |||||||||   | ||||| ||||||||
Db        182 AGGTAACAAAATCCCAAACATGTCTTCTTAGTCATGCAA 144


RESULT 13
CJ037346
LOCUS       CJ037346                662 bp    mRNA    linear   EST 22-OCT-2004
DEFINITION  CJ037346 full-length enriched swine cDNA library, adult testis Sus
            scrofa cDNA clone TES01G080111 5', mRNA sequence.
ACCESSION   CJ037346
VERSION     CJ037346.1  GI:54561241
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 662)
  AUTHORS   Uenishi,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H.,
            Okumura,N., Hamasima,N. and Awata,T.
  TITLE     PEDE (Pig EST Data Explorer): construction of a database for ESTs
            derived from porcine full-length cDNA libraries
  JOURNAL   Nucleic Acids Res. 32 (1), D484-D488 (2004)
   PUBMED   14681463
COMMENT     Contact: Hirohide Uenishi
            Animal Genome Laboratory, Genome Research Department
            National Institute of Agrobiological Sciences
            2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
            Tel: +81-29-838-8627
            Fax: +81-29-838-8627
            Email: huenishi@affrc.go.jp
            EST project with full-length enriched cDNA libraries carried out in
            Animal Genome Research Program (Japan) by National Institute of
            Agrobiological Sciences and STAFF-Institute
            Single pass sequencing of clones derived from oligo-capped cDNA
            library
            Vector sequences were eliminated by RepeatMasker version 2002/07/13
            and crossmatch version 0.990319
            Low quality bases were trimmed based on the quality values.
FEATURES             Location/Qualifiers
     source          1. .662
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /clone="TES01G080111"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /clone_lib="full-length enriched swine cDNA library, adult
                     testis"
ORIGIN

  Query Match             57.5%;  Score 23;  DB 7;  Length 662;
  Best Local Similarity   74.4%;  Pred. No. 2.6e+02;
  Matches   29;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||| ||  |||||||||||| | |  | ||| || |||
Db        611 AGGAAACTGACTCCCAAACAAGCTAAGTAACTTATCCAA 649


RESULT 14
CJ035837
LOCUS       CJ035837                790 bp    mRNA    linear   EST 22-OCT-2004
DEFINITION  CJ035837 full-length enriched swine cDNA library, adult testis Sus
            scrofa cDNA clone TES01E060090 5', mRNA sequence.
ACCESSION   CJ035837
VERSION     CJ035837.1  GI:54559732
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1  (bases 1 to 790)
  AUTHORS   Uenishi,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H.,
            Okumura,N., Hamasima,N. and Awata,T.
  TITLE     PEDE (Pig EST Data Explorer): construction of a database for ESTs
            derived from porcine full-length cDNA libraries
  JOURNAL   Nucleic Acids Res. 32 (1), D484-D488 (2004)
   PUBMED   14681463
COMMENT     Contact: Hirohide Uenishi
            Animal Genome Laboratory, Genome Research Department
            National Institute of Agrobiological Sciences
            2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
            Tel: +81-29-838-8627
            Fax: +81-29-838-8627
            Email: huenishi@affrc.go.jp
            EST project with full-length enriched cDNA libraries carried out in
            Animal Genome Research Program (Japan) by National Institute of
            Agrobiological Sciences and STAFF-Institute
            Single pass sequencing of clones derived from oligo-capped cDNA
            library
            Vector sequences were eliminated by RepeatMasker version 2002/07/13
            and crossmatch version 0.990319
            Low quality bases were trimmed based on the quality values.
FEATURES             Location/Qualifiers
     source          1. .790
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /clone="TES01E060090"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /clone_lib="full-length enriched swine cDNA library, adult
                     testis"
ORIGIN

  Query Match             57.5%;  Score 23;  DB 7;  Length 790;
  Best Local Similarity   74.4%;  Pred. No. 2.7e+02;
  Matches   29;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||| ||  |||||||||||| | |  | ||| || |||
Db        609 AGGAAACTGACTCCCAAACAAGCTAAGTAACTTATCCAA 647


RESULT 15
BG506830/c
LOCUS       BG506830                798 bp    mRNA    linear   EST 27-MAR-2001
DEFINITION  601861443F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4071154 5',
            mRNA sequence.
ACCESSION   BG506830
VERSION     BG506830.1  GI:13468347
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 798)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM916  row: e  column: 11
            High quality sequence stop: 2.
FEATURES             Location/Qualifiers
     source          1. .798
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4071154"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_77"
                     /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.9
                     kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             57.5%;  Score 23;  DB 2;  Length 798;
  Best Local Similarity   83.9%;  Pred. No. 2.7e+02;
  Matches   26;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         10 ACTCCCAAACAATCCATCTTACTCATGCAAC 40
              |||||||||| | || ||||| ||||||| |
Db        222 ACTCCCAAACCAGCCTTCTTAGTCATGCATC 192


Search completed: March  6, 2006, 01:57:49
Job time : 188.843 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 21:54:12 ; Search time 11.1834 Seconds
(without alignments)
6357.883 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278
Perfect score: 40
Sequence: 1 aggatacctactcccaaacaatccatcttactcatgcaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	40	100.0	1394	3	US-08-880-499-1	Sequence 1, Appli
	2	40	100.0	1394	3	US-08-880-499-2	Sequence 2, Appli
c	3	20.6	51.5	601	3	US-09-949-002-7790	Sequence 7790, Ap
c	4	20.6	51.5	601	3	US-09-949-002-7791	Sequence 7791, Ap
	5	20.6	51.5	1071	3	US-09-248-796A-5546	Sequence 5546, Ap
c	6	20.6	51.5	14882	3	US-09-949-002-782	Sequence 782, App
c	7	20.6	51.5	42075	3	US-09-949-016-14995	Sequence 14995, A
c	8	20.4	51.0	601	3	US-09-949-016-65770	Sequence 65770, A
	9	20.4	51.0	973	3	US-09-774-528-184	Sequence 184, App
	10	20.4	51.0	973	3	US-10-120-988-184	Sequence 184, App
	11	20.4	51.0	4201	3	US-08-945-056-4	Sequence 4, Appli
c	12	20.4	51.0	51770	3	US-09-949-016-13668	Sequence 13668, A
	13	20.4	51.0	84571	3	US-09-949-016-17420	Sequence 17420, A
c	14	20.4	51.0	126200	3	US-09-949-016-11824	Sequence 11824, A
c	15	20.4	51.0	126200	3	US-09-949-016-13193	Sequence 13193, A
	16	20.2	50.5	5099	3	US-09-610-040-5	Sequence 5, Appli
	17	20.2	50.5	5099	3	US-10-267-763-5	Sequence 5, Appli
	18	20	50.0	601	3	US-09-949-016-88655	Sequence 88655, A
	19	20	50.0	2615	3	US-09-620-312D-22	Sequence 22, Appl
	20	20	50.0	77586	3	US-09-949-016-13220	Sequence 13220, A
	21	20	50.0	77586	3	US-09-949-016-13221	Sequence 13221, A
	22	19.8	49.5	107	3	US-09-513-999C-19002	Sequence 19002, A
c	23	19.8	49.5	601	3	US-09-949-016-202698	Sequence 202698,
c	24	19.8	49.5	1017	3	US-09-710-279-2867	Sequence 2867, Ap
c	25	19.8	49.5	1164	3	US-09-134-001C-2534	Sequence 2534, Ap
c	26	19.8	49.5	2046	3	US-09-248-796A-6906	Sequence 6906, Ap
	27	19.8	49.5	2971	3	US-09-710-279-3449	Sequence 3449, Ap
c	28	19.8	49.5	9366	3	US-09-949-016-17437	Sequence 17437, A
	29	19.8	49.5	74096	3	US-09-949-016-11785	Sequence 11785, A
	30	19.8	49.5	74097	3	US-09-949-016-16239	Sequence 16239, A
c	31	19.8	49.5	265038	3	US-09-949-016-15779	Sequence 15779, A
	32	19.6	49.0	189	3	US-09-540-236-1562	Sequence 1562, Ap
	33	19.6	49.0	601	3	US-09-949-016-83110	Sequence 83110, A
	34	19.6	49.0	601	3	US-09-949-016-83111	Sequence 83111, A
c	35	19.6	49.0	601	3	US-09-949-002-4265	Sequence 4265, Ap
c	36	19.6	49.0	964	3	US-09-775-398-26	Sequence 26, Appl
	37	19.6	49.0	4322	2	US-08-537-342-1	Sequence 1, Appli
c	38	19.6	49.0	6330	3	US-09-949-016-15506	Sequence 15506, A
c	39	19.6	49.0	66986	3	US-09-596-002-29	Sequence 29, Appl
	40	19.6	49.0	75480	3	US-09-949-016-16090	Sequence 16090, A
c	41	19.6	49.0	80858	3	US-09-949-016-12659	Sequence 12659, A
c	42	19.6	49.0	80859	3	US-09-949-016-15715	Sequence 15715, A
	43	19.6	49.0	89892	3	US-09-949-016-13667	Sequence 13667, A
c	44	19.6	49.0	107937	3	US-09-949-016-17192	Sequence 17192, A
c	45	19.6	49.0	140315	3	US-09-949-016-14141	Sequence 14141, A

ALIGNMENTS

RESULT 1
US-08-880-499-1
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-880-499-1

Query Match 100.0%; Score 40; DB 3; Length 1394;

```
  Best Local Similarity   100.0%;  Pred. No. 6e-07;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 2
US-08-880-499-2
; Sequence 2, Application US/08880499
; Patent No. 6037523
;  GENERAL INFORMATION:
;    APPLICANT:  Albertson, Marc C.
;    APPLICANT:  Fox, Tim W.
;    APPLICANT:  Carl, Garnaat W.
;    APPLICANT:  Huffman, Gary A.
;    APPLICANT:  Kendall, Timmy L.
;    TITLE OF INVENTION:  MALE TISSUE-PREFERRED REGULATORY REGION
;    TITLE OF INVENTION:  AND METHOD OF USING SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PIONEER HI-BRED INTERNATIONAL, INC.
;      STREET:  Darwin Building, 7100 N.W. 62nd Ave., P.O.
;      STREET:  Box 1000
;      CITY:  Johnston
;      STATE:  Iowa
;      COUNTRY:  USA
;      ZIP:  50131
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/880,499
;      FILING DATE:  CONCURRENTLY HEREWITH
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sweeney, Patricia A.
;      REGISTRATION NUMBER:  32,733
;      REFERENCE/DOCKET NUMBER:  0578
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 248-4800
;      TELEFAX:  (515) 248-4844
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1394 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-880-499-2

  Query Match             100.0%;  Score 40;  DB 3;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 6e-07;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 3
US-09-949-002-7790/c
; Sequence 7790, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7790
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-7790

  Query Match             51.5%;  Score 20.6;  DB 3;  Length 601;
  Best Local Similarity   74.3%;  Pred. No. 51;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCA 38
              |||   ||  |||| ||||| ||| | ||||||||
Db        496 ATAAAAACAACCAATCAATCAATCATTCTCATGCA 462


RESULT 4
US-09-949-002-7791/c
; Sequence 7791, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7791
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-7791

  Query Match             51.5%;  Score 20.6;  DB 3;  Length 601;
  Best Local Similarity   74.3%;  Pred. No. 51;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCA 38
              |||   ||  |||| ||||| ||| | ||||||||
Db        379 ATAAAAACAACCAATCAATCAATCATTCTCATGCA 345


RESULT 5
US-09-248-796A-5546
; Sequence 5546, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5546
;   LENGTH: 1071
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-5546
```

```
  Query Match             51.5%;  Score 20.6;  DB 3;  Length 1071;
  Best Local Similarity   74.3%;  Pred. No. 57;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCA 38
              || || |||| ||||||   || |||||| || ||
Db        224 ATCCCCACTCTCAAACATGTCAACTTACTTATTCA 258


RESULT 6
US-09-949-002-782/c
; Sequence 782, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 782
;   LENGTH: 14882
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-782

  Query Match             51.5%;  Score 20.6;  DB 3;  Length 14882;
  Best Local Similarity   74.3%;  Pred. No. 99;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCA 38
              |||   ||  |||| ||||| ||| | ||||||||
Db       8503 ATAAAAACAACCAATCAATCAATCATTCTCATGCA 8469


RESULT 7
US-09-949-016-14995/c
; Sequence 14995, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14995
;   LENGTH: 42075
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14995

  Query Match             51.5%;  Score 20.6;  DB 3;  Length 42075;
  Best Local Similarity   85.2%;  Pred. No. 1.2e+02;
  Matches   23;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTA 30
              || ||||| | ||||| ||||||||||
Db      31591 ATTCCTACCCACAAACCATCCATCTTA 31565


RESULT 8
US-09-949-016-65770/c
; Sequence 65770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65770
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-65770

  Query Match             51.0%;  Score 20.4;  DB 3;  Length 601;
  Best Local Similarity   71.1%;  Pred. No. 61;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 GGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||| |||   | ||| ||||||  ||  |||| ||||
Db        414 GGATTCCTGGGCTCAAGCAATCCTCCTGCCTCAGGCAA 377


RESULT 9
US-09-774-528-184
; Sequence 184, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION:  No. 6743619el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE:  802
;  CURRENT APPLICATION NUMBER: US/09/774,528
;  CURRENT FILING DATE:  2001-01-30
;  NUMBER OF SEQ ID NOS:  441
;  SOFTWARE:  pt_FL_genes Version 2.0
; SEQ ID NO 184
;   LENGTH: 973
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (231)..(749)
US-09-774-528-184

  Query Match             51.0%;  Score 20.4;  DB 3;  Length 973;
  Best Local Similarity   71.1%;  Pred. No. 68;
```

```
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 GGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              ||| | | ||||| ||| || || ||  |  |||||||
Db        626 GGAAATCAACTCCAAAAGAAACCCTCGAAACCATGCAA 663


RESULT 10
US-10-120-988-184
; Sequence 184, Application US/10120988
; Patent No. 6919193
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6919193el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 802CON
;  CURRENT APPLICATION NUMBER: US/10/120,988
;  CURRENT FILING DATE:  2002-04-11
;  PRIOR APPLICATION NUMBER: 09/774,528
;  PRIOR FILING DATE: 2001-01-30
;  NUMBER OF SEQ ID NOS: 441
;  SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 184
;   LENGTH: 973
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (231)..(749)
US-10-120-988-184

  Query Match             51.0%;  Score 20.4;  DB 3;  Length 973;
  Best Local Similarity   71.1%;  Pred. No. 68;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 GGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              ||| | | ||||| ||| || || ||  |  |||||||
Db        626 GGAAATCAACTCCAAAAGAAACCCTCGAAACCATGCAA 663


RESULT 11
US-08-945-056-4
; Sequence 4, Application US/08945056
; Patent No. 6077994
;  GENERAL INFORMATION:
;    APPLICANT:  Coupland, George M.
;    APPLICANT:  Putterill, Joanna J.
;    TITLE OF INVENTION:  Genetic control of flowering
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon & Vanderhye PC
;      STREET:  8th Floor, 1100 No. 6077994th Glebe Road
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  United States of America
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/945,056
;      FILING DATE:  20-OCT-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02561
;      FILING DATE:  01-NOV-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9422083.7
;      FILING DATE:  02-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ms Mary J Wilson
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  620-17
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816-4000
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4201 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Arabidopsis thaliana
;      STRAIN:  Landsberg erecta
;    POSITION IN GENOME:
;      MAP POSITION:  chromosome 5
US-08-945-056-4

  Query Match             51.0%;  Score 20.4;  DB 3;  Length 4201;
  Best Local Similarity   80.0%;  Pred. No. 92;
  Matches   24;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 GATACCTACTCCCAAACAATCCATCTTACT 32
              ||||||  |||||| || ||| | ||||||
Db       2631 GATACCAGCTCCCACACCATCAAACTTACT 2660


RESULT 12
US-09-949-016-13668/c
; Sequence 13668, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13668
;   LENGTH: 51770
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(51770)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13668

  Query Match             51.0%;  Score 20.4;  DB 3;  Length 51770;
  Best Local Similarity   71.1%;  Pred. No. 1.6e+02;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 GGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||| |||   | ||| ||||||  ||  |||| ||||
Db       9851 GGATTCCTGGGCTCAAGCAATCCTCCTGCCTCAGGCAA 9814


RESULT 13
US-09-949-016-17420
; Sequence 17420, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17420
;   LENGTH: 84571
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(84571)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17420

  Query Match           51.0%;  Score 20.4;  DB 3;  Length 84571;
  Best Local Similarity   71.1%;  Pred. No. 1.7e+02;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 GGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              ||| | | ||||| ||| || || ||  |  |||||||
Db      24559 GGAAATCAACTCCAAAAGAAACCTTCAAAACCATGCAA 24596


RESULT 14
US-09-949-016-11824/c
; Sequence 11824, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11824
;   LENGTH: 126200
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(126200)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11824

  Query Match           51.0%;  Score 20.4;  DB 3;  Length 126200;
  Best Local Similarity   71.1%;  Pred. No. 1.9e+02;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 GGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              ||| | | ||||| |||  | || ||  | ||||||||
Db      53133 GGAAATCAACTCCAAAAGGAACCTTCAAAATCATGCAA 53096


RESULT 15
US-09-949-016-13193/c
; Sequence 13193, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13193
;   LENGTH: 126200
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(126200)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13193

  Query Match           51.0%;  Score 20.4;  DB 3;  Length 126200;
  Best Local Similarity   71.1%;  Pred. No. 1.9e+02;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 GGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              ||| | | ||||| |||  | || ||  | ||||||||
Db      53133 GGAAATCAACTCCAAAAGGAACCTTCAAAATCATGCAA 53096


Search completed: March  5, 2006, 22:36:28
Job time : 11.1834 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 23:29:31 ; Search time 57.5163 Seconds
(without alignments)
5750.985 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278
Perfect score: 40
Sequence: 1 aggatacctactcccaaacaatccatcttactcatgcaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	100.0	40	8	US-10-713-381-6	Sequence 6, Appli
2	40	100.0	50	8	US-10-713-381-5	Sequence 5, Appli
3	40	100.0	158	8	US-10-713-381-3	Sequence 3, Appli
4	40	100.0	255	8	US-10-713-381-9	Sequence 9, Appli
5	40	100.0	1394	8	US-10-713-381-1	Sequence 1, Appli
6	40	100.0	1394	8	US-10-713-381-2	Sequence 2, Appli
c 7	24	60.0	85779	8	US-10-719-993-6880	Sequence 6880, Ap
8	23.2	58.0	706	7	US-10-424-599-31006	Sequence 31006, A
c 9	23	57.5	636	4	US-09-925-065A-670302	Sequence 670302,
10	22.6	56.5	294	8	US-10-425-115-143168	Sequence 143168,
11	22.4	56.0	315	7	US-10-424-599-83179	Sequence 83179, A
12	22.4	56.0	226215	5	US-10-087-192-1948	Sequence 1948, Ap
c 13	22.2	55.5	611	4	US-09-925-065A-591011	Sequence 591011,
c 14	22.2	55.5	611	4	US-09-925-065A-591012	Sequence 591012,
c 15	22.2	55.5	611	4	US-09-925-065A-591013	Sequence 591013,
16	22.2	55.5	613	4	US-09-925-065A-253727	Sequence 253727,
c 17	22	55.0	628	4	US-09-925-065A-888321	Sequence 888321,
c 18	21.6	54.0	598	4	US-09-925-065A-393490	Sequence 393490,
c 19	21.6	54.0	4110	8	US-10-473-126-342	Sequence 342, App
c 20	21.6	54.0	6309	6	US-10-311-455-277	Sequence 277, App
21	21.4	53.5	560	4	US-09-925-065A-281837	Sequence 281837,
22	21.4	53.5	560	4	US-09-925-065A-281838	Sequence 281838,
c 23	21.4	53.5	620	4	US-09-925-065A-189248	Sequence 189248,
24	21.4	53.5	653	4	US-09-925-065A-63167	Sequence 63167, A
c 25	21.4	53.5	1341	3	US-09-938-842A-1555	Sequence 1555, Ap
c 26	21.4	53.5	1341	3	US-09-938-842A-1555	Sequence 1555, Ap
c 27	21.4	53.5	6681	6	US-10-311-455-128	Sequence 128, App
c 28	21.4	53.5	6681	6	US-10-240-452-4	Sequence 4, Appli
c 29	21.4	53.5	8168	6	US-10-311-455-901	Sequence 901, App
c 30	21.4	53.5	8168	7	US-10-240-454-23	Sequence 23, Appl
c 31	21.4	53.5	24401	10	US-11-097-143-2335	Sequence 2335, Ap
c 32	21.4	53.5	3673778	6	US-10-312-841-1	Sequence 1, Appli
c 33	21.4	53.5	3673778	6	US-10-312-841-2	Sequence 2, Appli
c 34	21	52.5	293	7	US-10-424-599-24043	Sequence 24043, A
35	21	52.5	617	4	US-09-925-065A-285326	Sequence 285326,
36	21	52.5	617	4	US-09-925-065A-285327	Sequence 285327,
37	21	52.5	617	4	US-09-925-065A-285328	Sequence 285328,
c 38	21	52.5	653	4	US-09-925-065A-937247	Sequence 937247,
c 39	21	52.5	931	7	US-10-424-599-20320	Sequence 20320, A
40	21	52.5	2000	7	US-10-260-238-2594	Sequence 2594, Ap
c 41	21	52.5	2792	7	US-10-041-018-361	Sequence 361, App
c 42	21	52.5	3117	7	US-10-041-018-362	Sequence 362, App
c 43	21	52.5	5507	8	US-10-473-126-198	Sequence 198, App
c 44	21	52.5	5507	8	US-10-473-126-344	Sequence 344, App
c 45	21	52.5	8404	6	US-10-311-455-1568	Sequence 1568, Ap

ALIGNMENTS

RESULT 1
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6

Query Match 100.0%; Score 40; DB 8; Length 40;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
||||||||||||||||||||||||||||||||||||||||
Db 1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40

RESULT 2
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME

```
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-5

  Query Match             100.0%;  Score 40;  DB 8;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db         11 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 50


RESULT 3
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 158
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-3

  Query Match             100.0%;  Score 40;  DB 8;  Length 158;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db         86 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 125


RESULT 4
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 255
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-9

  Query Match             100.0%;  Score 40;  DB 8;  Length 255;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db         99 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 138


RESULT 5
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1394
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-1

  Query Match             100.0%;  Score 40;  DB 8;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 8e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 6
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1394
;   TYPE: DNA
```

```
;   ORGANISM: Zea mays
US-10-713-381-2

  Query Match           100.0%;  Score 40;  DB 8;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 8e-06;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||||||||||||||||||||||||||||||||||||||
Db       1239 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 1278


RESULT 7
US-10-719-993-6880/c
; Sequence 6880, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001496
;  CURRENT APPLICATION NUMBER: US/10/719,993
;  CURRENT FILING DATE:  2003-11-24
;  NUMBER OF SEQ ID NOS: 55342
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6880
;   LENGTH: 85779
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(85779)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6880

  Query Match            60.0%;  Score 24;  DB 8;  Length 85779;
  Best Local Similarity  84.4%;  Pred. No. 47;
  Matches   27;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCAT 35
              | || ||||||||||||||| || | ||||||
Db      16082 AAACTTACTCCCAAACAATCTATGTCACTCAT 16051


RESULT 8
US-10-424-599-31006
; Sequence 31006, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 31006
;   LENGTH: 706
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_128002C.1
US-10-424-599-31006

  Query Match            58.0%;  Score 23.2;  DB 7;  Length 706;
  Best Local Similarity  77.8%;  Pred. No. 39;
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          5 TACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              | ||| |||||||||  ||||||   ||||| ||||
Db        455 TCCCTTCTCCCAAACCCTCCATCAACCTCATACAAC 490


RESULT 9
US-09-925-065A-670302/c
; Sequence 670302, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 670302
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-670302

  Query Match            57.5%;  Score 23;  DB 4;  Length 636;
  Best Local Similarity  74.4%;  Pred. No. 45;
  Matches   29;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||| ||  | |  ||||| ||||| ||| |||| ||||
Db        474 AGGAAACGGATTAACAAACTATCCACCTTTCTCAAGCAA 436


RESULT 10
US-10-425-115-143168
; Sequence 143168, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 143168
;   LENGTH: 294
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_62049C.1
US-10-425-115-143168

  Query Match            56.5%;  Score 22.6;  DB 8;  Length 294;
  Best Local Similarity  86.2%;  Pred. No. 56;
  Matches   25;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACT 32
              ||||||  ||||| ||| ||||||||||
Db        226 ATACCTGGTCCCATACACTCCATCTTACT 254
```

```
RESULT 11
US-10-424-599-83179
; Sequence 83179, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 83179
;   LENGTH: 315
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_46127C.1
US-10-424-599-83179

  Query Match             56.0%;  Score 22.4;  DB 7;  Length 315;
  Best Local Similarity   81.2%;  Pred. No. 69;
  Matches   26;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCAT 35
              | ||  ||||| ||||||||||||||||| || |
Db        211 AAACGAACTCCTAAACAATCCATCTTAATCCT 242


RESULT 12
US-10-087-192-1948
; Sequence 1948, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1948
;   LENGTH: 226215
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-087-192-1948

  Query Match             56.0%;  Score 22.4;  DB 5;  Length 226215;
  Best Local Similarity   72.5%;  Pred. No. 2.5e+02;
  Matches   29;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              || || ||| || |||||||   |||||||||| |   ||
Db     183206 AGAATTCCTGCTTCCAAACACAGCATCTTACTCCTTTTAC 183245


RESULT 13
US-09-925-065A-591011/c
; Sequence 591011, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591011
;   LENGTH: 611
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-591011

  Query Match             55.5%;  Score 22.2;  DB 4;  Length 611;
  Best Local Similarity   88.9%;  Pred. No. 94;
  Matches   24;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         14 CCAAACAATCCATCTTACTCATGCAAC 40
              |||||||| ||| ||||||||||||
Db        351 CCAAACAAACCACCTTACTCATGCAGC 325


RESULT 14
US-09-925-065A-591012/c
; Sequence 591012, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591012
;   LENGTH: 611
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-591012

  Query Match             55.5%;  Score 22.2;  DB 4;  Length 611;
  Best Local Similarity   88.9%;  Pred. No. 94;
  Matches   24;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         14 CCAAACAATCCATCTTACTCATGCAAC 40
              |||||||| ||| ||||||||||||
Db        351 CCAAACAAACCACCTTACTCATGCAGC 325


RESULT 15
US-09-925-065A-591013/c
; Sequence 591013, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591013
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-591013

Query Match 55.5%; Score 22.2; DB 4; Length 611;
Best Local Similarity 88.9%; Pred. No. 94;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         14 CCAAACAATCCATCTTACTCATGCAAC 40
              |||||||| ||| ||||||||||||| |
Db        351 CCAAACAAACCACCTTACTCATGCAGC 325
```

Search completed: March 6, 2006, 03:55:19
Job time : 59.5163 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 01:58:03 ; Search time 41.7555 Seconds
(without alignments)
2100.307 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278
Perfect score: 40
Sequence: 1 aggatacctactcccaaacaatccatcttactcatgcaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	23	57.5	636	6	US-09-925-065A-670302	Sequence 670302,
	2	22.6	56.5	3242	8	US-10-750-185-58624	Sequence 58624, A
	3	22.6	56.5	3242	8	US-10-750-623-58624	Sequence 58624, A
c	4	22.2	55.5	611	6	US-09-925-065A-591011	Sequence 591011,
c	5	22.2	55.5	611	6	US-09-925-065A-591012	Sequence 591012,
c	6	22.2	55.5	611	6	US-09-925-065A-591013	Sequence 591013,
	7	22.2	55.5	613	6	US-09-925-065A-253727	Sequence 253727,
c	8	22	55.0	628	6	US-09-925-065A-888321	Sequence 888321,
c	9	21.6	54.0	598	6	US-09-925-065A-393490	Sequence 393490,
	10	21.4	53.5	560	6	US-09-925-065A-281837	Sequence 281837,
	11	21.4	53.5	560	6	US-09-925-065A-281838	Sequence 281838,
c	12	21.4	53.5	620	6	US-09-925-065A-189248	Sequence 189248,
	13	21.4	53.5	653	6	US-09-925-065A-63167	Sequence 63167, A
	14	21	52.5	201	12	US-11-124-367A-24503	Sequence 24503, A
	15	21	52.5	617	6	US-09-925-065A-285326	Sequence 285326,
	16	21	52.5	617	6	US-09-925-065A-285327	Sequence 285327,
	17	21	52.5	617	6	US-09-925-065A-285328	Sequence 285328,
c	18	21	52.5	653	6	US-09-925-065A-937247	Sequence 937247,
c	19	21	52.5	1718	8	US-10-750-185-62555	Sequence 62555, A
c	20	21	52.5	1718	8	US-10-750-623-62555	Sequence 62555, A
	21	20.8	52.0	1109	8	US-10-750-185-49907	Sequence 49907, A
	22	20.8	52.0	1109	8	US-10-750-623-49907	Sequence 49907, A
	23	20.8	52.0	1217	8	US-10-750-185-57406	Sequence 57406, A
	24	20.8	52.0	1217	8	US-10-750-623-57406	Sequence 57406, A
c	25	20.6	51.5	50	12	US-11-175-859-103341	Sequence 103341,
c	26	20.6	51.5	569	6	US-09-925-065A-127036	Sequence 127036,
	27	20.6	51.5	586	6	US-09-925-065A-610819	Sequence 610819,
c	28	20.6	51.5	664	6	US-09-925-065A-696032	Sequence 696032,
c	29	20.6	51.5	664	6	US-09-925-065A-696033	Sequence 696033,
c	30	20.4	51.0	201	8	US-10-995-561-70646	Sequence 70646, A
c	31	20.4	51.0	413	6	US-09-925-065A-360599	Sequence 360599,
c	32	20.4	51.0	470	6	US-09-925-065A-951077	Sequence 951077,
c	33	20.4	51.0	507	6	US-09-925-065A-589111	Sequence 589111,
c	34	20.4	51.0	507	6	US-09-925-065A-589112	Sequence 589112,
c	35	20.4	51.0	507	6	US-09-925-065A-589113	Sequence 589113,
c	36	20.4	51.0	507	6	US-09-925-065A-589114	Sequence 589114,
	37	20.4	51.0	519	6	US-09-925-065A-166725	Sequence 166725,
c	38	20.4	51.0	545	6	US-09-925-065A-241631	Sequence 241631,
	39	20.4	51.0	552	6	US-09-925-065A-785033	Sequence 785033,
	40	20.4	51.0	558	6	US-09-925-065A-659195	Sequence 659195,
	41	20.4	51.0	566	6	US-09-925-065A-910942	Sequence 910942,
	42	20.4	51.0	568	6	US-09-925-065A-104305	Sequence 104305,
	43	20.4	51.0	576	6	US-09-925-065A-573143	Sequence 573143,
c	44	20.4	51.0	583	6	US-09-925-065A-575922	Sequence 575922,
c	45	20.4	51.0	583	6	US-09-925-065A-575923	Sequence 575923,

ALIGNMENTS

RESULT 1
US-09-925-065A-670302/c
; Sequence 670302, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 670302
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-670302

Query Match 57.5%; Score 23; DB 6; Length 636;
Best Local Similarity 74.4%; Pred. No. 11;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||| ||  | |  ||||| ||||| ||| |||| ||||
Db        474 AGGAAACGGATTAACAAACTATCCACCTTTCTCAAGCAA 436
```

RESULT 2
US-10-750-185-58624
; Sequence 58624, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:

```
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 58624
;   LENGTH: 3242
;   TYPE: DNA
;   ORGANISM: Bovine  19866880685452
US-10-750-185-58624

  Query Match             56.5%;  Score 22.6;  DB 8;  Length 3242;
  Best Local Similarity   75.7%;  Pred. No. 23;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||| || |||| || || | |||| | |||||||
Db       2052 ATACCACCTACCAACCAGTCAAGCTTATTGATGCAAC 2088


RESULT 3
US-10-750-623-58624
; Sequence 58624, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 58624
;   LENGTH: 3242
;   TYPE: DNA
;   ORGANISM: Bovine  19866880685452
US-10-750-623-58624

  Query Match             56.5%;  Score 22.6;  DB 8;  Length 3242;
  Best Local Similarity   75.7%;  Pred. No. 23;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              ||||| || |||| || || | |||| | |||||||
Db       2052 ATACCACCTACCAACCAGTCAAGCTTATTGATGCAAC 2088


RESULT 4
US-09-925-065A-591011/c
; Sequence 591011, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591011
;   LENGTH: 611
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-591011

  Query Match             55.5%;  Score 22.2;  DB 6;  Length 611;
  Best Local Similarity   88.9%;  Pred. No. 22;
  Matches   24;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         14 CCAAACAATCCATCTTACTCATGCAAC 40
              |||||||| ||| |||||||||||| |
Db        351 CCAAACAAACCACCTTACTCATGCAGC 325


RESULT 5
US-09-925-065A-591012/c
; Sequence 591012, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591012
;   LENGTH: 611
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-591012

  Query Match             55.5%;  Score 22.2;  DB 6;  Length 611;
  Best Local Similarity   88.9%;  Pred. No. 22;
  Matches   24;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         14 CCAAACAATCCATCTTACTCATGCAAC 40
              |||||||| ||| |||||||||||| |
Db        351 CCAAACAAACCACCTTACTCATGCAGC 325


RESULT 6
US-09-925-065A-591013/c
; Sequence 591013, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 591013
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-591013

Query Match 55.5%; Score 22.2; DB 6; Length 611;
Best Local Similarity 88.9%; Pred. No. 22;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 CCAAACAATCCATCTTACTCATGCAAC 40
||||||||| ||| ||||||||||||| |
Db 351 CCAAACAAACCACCTTACTCATGCAGC 325

RESULT 7
US-09-925-065A-253727
; Sequence 253727, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 253727
; LENGTH: 613
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-253727

Query Match 55.5%; Score 22.2; DB 6; Length 613;
Best Local Similarity 73.0%; Pred. No. 22;
Matches 27; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 2 GGATACCTACTCCCAAACAATCCATCTTACTCATGCA 38
||| | | ||||| |||:|| || || | |||||||
Db 292 GGAAATCAACTCCAAAASAAACCCTCAAAATCATGCA 328

RESULT 8
US-09-925-065A-888321/c
; Sequence 888321, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 888321
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-888321

Query Match 55.0%; Score 22; DB 6; Length 628;
Best Local Similarity 73.7%; Pred. No. 27;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 2 GGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
||| | | ||||| ||| |||||||| | |||| ||
Db 224 GGAAATCAACTCCAAAAGAATCCATCAAAACCATGGAA 187

RESULT 9
US-09-925-065A-393490/c
; Sequence 393490, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 393490
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-393490

Query Match 54.0%; Score 21.6; DB 6; Length 598;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 TACCTACTCCCAAACAATCCATCTTACT 32
|| || ||| ||| ||||||||||||||
Db 228 TATCTSCTCTCAATCAATCCATCTTACT 201

```
RESULT 10
US-09-925-065A-281837
; Sequence 281837, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 281837
;   LENGTH: 560
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-281837

  Query Match             53.5%;  Score 21.4;  DB 6;  Length 560;
  Best Local Similarity   71.8%;  Pred. No. 46;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||| | | ||||| |||  | || ||  || |||||||
Db        373 AGGAAATCAACTCCAAAAGGAACCCTCGAACCCATGCAA 411


RESULT 11
US-09-925-065A-281838
; Sequence 281838, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 281838
;   LENGTH: 560
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-281838

  Query Match             53.5%;  Score 21.4;  DB 6;  Length 560;
  Best Local Similarity   71.8%;  Pred. No. 46;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||| | | ||||| |||  | || ||  || |||||||
Db        373 AGGAAATCAACTCCAAAAGGAACCCTCGAACCCATGCAA 411


RESULT 12
US-09-925-065A-189248/c
; Sequence 189248, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 189248
;   LENGTH: 620
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-189248

  Query Match             53.5%;  Score 21.4;  DB 6;  Length 620;
  Best Local Similarity   75.8%;  Pred. No. 47;
  Matches   25;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTC 33
              | ||||:||||| || | | || ||||| ||||
Db        175 ATGATAMCTACTTCCTATCTATGCATCTCACTC 143


RESULT 13
US-09-925-065A-63167
; Sequence 63167, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63167
;   LENGTH: 653
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-63167

  Query Match             53.5%;  Score 21.4;  DB 6;  Length 653;
  Best Local Similarity   71.8%;  Pred. No. 47;
```

Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||| |||||| |||  |||   || || |||  |||||
Db        186 AGGAAACCTACACCCTGACATGTCAACTAACTACTGCAA 224
```

RESULT 14
US-11-124-367A-24503
; Sequence 24503, Application US/11124367A
; Publication No. US20060024700A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: Hongjin Huang
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Fibrosis Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001519.ORD
; CURRENT APPLICATION NUMBER: US/11/124,367A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,846
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/582,609
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US 60/599,554
; PRIOR FILING DATE: 2004-08-09
; NUMBER OF SEQ ID NOS: 34460
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24503
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-367A-24503

Query Match 52.5%; Score 21; DB 12; Length 201;
Best Local Similarity 73.0%; Pred. No. 52;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          3 GATACCTACTCCCAAACAATCCATCTTACTCATGCAA 39
              |||||||  || |||| ||| ||| |||  |||  ||
Db         43 GATACCTTGTCTCAAAAAATACATATTAAACATTAAA 79
```

RESULT 15
US-09-925-065A-285326
; Sequence 285326, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 285326
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-285326

Query Match 52.5%; Score 21; DB 6; Length 617;
Best Local Similarity 73.0%; Pred. No. 68;

Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          4 ATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40
              |||| |||| ||||||||||||   |  |   |||||
Db        387 ATACATACTTCCAAACAATCCAGAGTTTTGTAGCAAC 423
```

Search completed: March 6, 2006, 10:32:52
Job time : 42.7555 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:38:03 ; Search time 685.165 Seconds
(without alignments)
13025.213 Million cell updates/sec

Title: US-10-713-381-1_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match %	Length	DB	ID	Description
	1	157	100.0	255	6	AX224402	AX224402 Sequence
	2	157	100.0	1394	6	BD062176	BD062176 Male tiss
	3	157	100.0	1394	6	BD062177	BD062177 Male tiss
	4	157	100.0	1394	6	AX224394	AX224394 Sequence
	5	157	100.0	1394	6	AX224395	AX224395 Sequence
	6	157	100.0	3343	15	AF360356	AF360356 Zea mays
	7	146	93.0	158	6	AX224396	AX224396 Sequence
c	8	54	34.4	110000	15	AP008209_086	Continuation (87 o
c	9	54	34.4	137327	15	AC135206	AC135206 Oryza sat
	10	50	31.8	50	6	AX224398	AX224398 Sequence
	11	40	25.5	40	6	AX224399	AX224399 Sequence
c	12	34.8	22.2	171896	5	BX950854	BX950854 Zebrafish
	13	34.6	22.0	110127	8	AC104134	AC104134 Homo sapi
	14	34.4	21.9	214946	8	AC136896	AC136896 Homo sapi
	15	33.8	21.5	141554	14	CR847532	CR847532 Danio rer
	16	33.8	21.5	214178	5	BX640577	BX640577 Zebrafish
	17	33.4	21.3	211190	14	AC152229	AC152229 Bos tauru
c	18	33.4	21.3	237262	14	AC162712	AC162712 Bos tauru
c	19	33.2	21.1	117764	8	AC011290	AC011290 Homo sapi
c	20	33.2	21.1	162481	14	AC091864	AC091864 Homo sapi
	21	33.2	21.1	177594	8	AC146134	AC146134 Pan trogl
	22	32.8	20.9	82642	8	AC103745	AC103745 Homo sapi
c	23	32.8	20.9	93610	8	AC006602	AC006602 Homo sapi
	24	32.8	20.9	130031	14	AC135264	AC135264 Homo sapi
c	25	32.8	20.9	148598	8	HSBA51C14	AL121875 Human DNA
c	26	32.8	20.9	158316	8	AC135350	AC135350 Homo sapi
	27	32.8	20.9	169771	8	AC087485	AC087485 Homo sapi
	28	32.8	20.9	169867	14	AC019229	AC019229 Homo sapi
c	29	32.8	20.9	193569	8	AC149044	AC149044 Pan trogl
	30	32.6	20.8	53688	14	AC099973	AC099973 Mus muscu
	31	32.6	20.8	110000	14	AC147308_0	AC147308 Pan trogl
	32	32.6	20.8	146080	5	AL953855	AL953855 Zebrafish
	33	32.6	20.8	165845	14	AC127012	AC127012 Rattus no
c	34	32.6	20.8	204190	14	AC140696	AC140696 Rattus no
c	35	32.6	20.8	252534	14	AC097405	AC097405 Rattus no
	36	32.4	20.6	2070	9	BC020027	BC020027 Mus muscu
	37	32.4	20.6	2096	9	BC038066	BC038066 Mus muscu
	38	32.4	20.6	2433	9	BC055469	BC055469 Mus muscu
	39	32.4	20.6	3092	9	BC094506	BC094506 Mus muscu
c	40	32.4	20.6	110000	1	AE017197_05	Continuation (6 of
	41	32.4	20.6	134895	9	AC159811	AC159811 Mus muscu
	42	32.4	20.6	168889	9	AL731779	AL731779 Mouse DNA
c	43	32.4	20.6	182847	14	AC148229	AC148229 Gallus ga
	44	32.4	20.6	221338	5	AC147437	AC147437 Gallus ga
c	45	32.4	20.6	253762	9	AC157998	AC157998 Mus muscu

ALIGNMENTS

RESULT 1
AX224402
LOCUS AX224402 255 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 9 from Patent WO0160997.
ACCESSION AX224402
VERSION AX224402.1 GI:15554644
KEYWORDS .
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 9 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES Location/Qualifiers
source 1. .255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN

Query Match 100.0%; Score 157; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 5.6e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 74

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 134

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db        135 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171
```

```
RESULT 2
BD062176
LOCUS       BD062176               1394 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Male tissue-preferred regulatory region and method of using same.
ACCESSION   BD062176
VERSION     BD062176.1  GI:22607781
KEYWORDS    JP 2001520523-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1394)
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
            Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: JP 2001520523-A 1 30-OCT-2001;
            PIONEER HI BRED INTERNATIONAL INC
COMMENT     PN   JP 2001520523-A/1
            PD   30-OCT-2001
            PF   19-JUN-1998 JP 1999504910
            PR   23-JUN-1997 US    08/880499
            PI   MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
            PI   TIMMY L KENDALL
            PC   C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
            ,C07K14/34,C12Q1/68,
            PC   A01H5/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES                Location/Qualifiers
     source             1. .1394
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 157;  DB 6;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-40;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 3
BD062177
LOCUS       BD062177               1394 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Male tissue-preferred regulatory region and method of using same.
ACCESSION   BD062177
VERSION     BD062177.1  GI:22607782
KEYWORDS    JP 2001520523-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1394)
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
            Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: JP 2001520523-A 2 30-OCT-2001;
            PIONEER HI BRED INTERNATIONAL INC
COMMENT     PN   JP 2001520523-A/2
            PD   30-OCT-2001
            PF   19-JUN-1998 JP 1999504910
            PR   23-JUN-1997 US    08/880499
            PI   MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
            PI   TIMMY L KENDALL
            PC   C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
            ,C07K14/34,C12Q1/68,
            PC   A01H5/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES                Location/Qualifiers
     source             1. .1394
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 157;  DB 6;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-40;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 4
AX224394
LOCUS       AX224394               1394 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 1 from Patent WO0160997.
ACCESSION   AX224394
VERSION     AX224394.1  GI:15554636
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 1 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES                Location/Qualifiers
     source             1. .1394
                        /organism="Zea mays"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:4577"
ORIGIN

  Query Match             100.0%;  Score 157;  DB 6;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-40;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
```

```
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 5
AX224395
LOCUS       AX224395               1394 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 2 from Patent WO0160997.
ACCESSION   AX224395
VERSION     AX224395.1  GI:15554637
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 2 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1394
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             100.0%;  Score 157;  DB 6;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-40;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 6
AF360356
LOCUS       AF360356               3343 bp    DNA     linear   PLN 12-MAY-2001
DEFINITION  Zea mays male fertility protein (Ms45) gene, complete cds.
ACCESSION   AF360356
VERSION     AF360356.1  GI:14028756
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3343)
  AUTHORS   Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
  TITLE     Cloning of Ms45, a gene required for male ferility from Zea mays
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3343)
  AUTHORS   Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
            Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
            IA 50131-1004, USA
FEATURES             Location/Qualifiers
     source          1. .3343
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /chromosome="9L"
     gene            <1392. .>2942
                     /gene="Ms45"
     mRNA            join(<1392. .1768,1898. .2182,2280. .2447,2534. .>2942)
                     /gene="Ms45"
                     /product="male fertility protein"
     CDS             join(1392. .1768,1898. .2182,2280. .2447,2534. .2942)
                     /gene="Ms45"
                     /codon_start=1
                     /product="male fertility protein"
                     /protein_id="AAK52489.1"
                     /db_xref="GI:14028757"
                     /translation="MEKRNLQWRRGRDGIVQYPHLFFAALALALLVADPFGLSPLAEV
                     DYRPVKHELAPYGEVMGSWPRDNASRLRRGRLEFVGEVFGPESIEFDLQGRGPYAGLA
                     DGRVVRWMGEEAGWETFAVMNPDWSEEVCANGVNSTTRKQHEKEEFCGRPLGLRFHGE
                     TGELYVADAYYGLMVVGQSGGVASSVAREADGDPIRFANDLDVHRNGSVFFTDTSMRY
                     SRKDHLNILLEGEGTGRLLRYDPETSGVHVVLKGLVFPNGVQISEDHQFLLFSETTNC
                     RIMRYWLEGPRAGEVEVFANLPGFPDNVRSNGRGQFWVAIDCCRTPAQEVFAKRPWLR
                     TLYFKFPLSLKVLTWKAARRMHTVLALLDGEGRVVEVLEDRGHEVMKLVSEVREVGRK
                     LWIGTVAHNHIATIPYPLED"
ORIGIN

  Query Match             100.0%;  Score 157;  DB 15;  Length 3343;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-40;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 7
AX224396
LOCUS       AX224396               158 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 3 from Patent WO0160997.
ACCESSION   AX224396
VERSION     AX224396.1  GI:15554638
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 3 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .158
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             93.0%;  Score 146;  DB 6;  Length 158;
  Best Local Similarity   99.4%;  Pred. No. 2.1e-36;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 59
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy         60 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 119
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120
```

```
Qy        120 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||||
Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158
```

RESULT 8
AP008209_086/c
WPCOMMENT
Sequence split into 362 fragments LOCUS AP008209 Accession AP008209

Fragment Name	Begin	End
AP008209_000	1	110000
AP008209_001	100001	210000
AP008209_002	200001	310000
AP008209_003	300001	410000
AP008209_004	400001	510000
AP008209_005	500001	610000
AP008209_006	600001	710000
AP008209_007	700001	810000
AP008209_008	800001	910000
AP008209_009	900001	1010000
AP008209_010	1000001	1110000
AP008209_011	1100001	1210000
AP008209_012	1200001	1310000
AP008209_013	1300001	1410000
AP008209_014	1400001	1510000
AP008209_015	1500001	1610000
AP008209_016	1600001	1710000
AP008209_017	1700001	1810000
AP008209_018	1800001	1910000
AP008209_019	1900001	2010000
AP008209_020	2000001	2110000
AP008209_021	2100001	2210000
AP008209_022	2200001	2310000
AP008209_023	2300001	2410000
AP008209_024	2400001	2510000
AP008209_025	2500001	2610000
AP008209_026	2600001	2710000
AP008209_027	2700001	2810000
AP008209_028	2800001	2910000
AP008209_029	2900001	3010000
AP008209_030	3000001	3110000
AP008209_031	3100001	3210000
AP008209_032	3200001	3310000
AP008209_033	3300001	3410000
AP008209_034	3400001	3510000
AP008209_035	3500001	3610000
AP008209_036	3600001	3710000
AP008209_037	3700001	3810000
AP008209_038	3800001	3910000
AP008209_039	3900001	4010000
AP008209_040	4000001	4110000
AP008209_041	4100001	4210000
AP008209_042	4200001	4310000
AP008209_043	4300001	4410000
AP008209_044	4400001	4510000
AP008209_045	4500001	4610000
AP008209_046	4600001	4710000
AP008209_047	4700001	4810000
AP008209_048	4800001	4910000
AP008209_049	4900001	5010000
AP008209_050	5000001	5110000
AP008209_051	5100001	5210000
AP008209_052	5200001	5310000
AP008209_053	5300001	5410000
AP008209_054	5400001	5510000
AP008209_055	5500001	5610000
AP008209_056	5600001	5710000
AP008209_057	5700001	5810000
AP008209_058	5800001	5910000
AP008209_059	5900001	6010000
AP008209_060	6000001	6110000
AP008209_061	6100001	6210000
AP008209_062	6200001	6310000
AP008209_063	6300001	6410000
AP008209_064	6400001	6510000
AP008209_065	6500001	6610000
AP008209_066	6600001	6710000
AP008209_067	6700001	6810000
AP008209_068	6800001	6910000
AP008209_069	6900001	7010000
AP008209_070	7000001	7110000
AP008209_071	7100001	7210000
AP008209_072	7200001	7310000
AP008209_073	7300001	7410000
AP008209_074	7400001	7510000
AP008209_075	7500001	7610000
AP008209_076	7600001	7710000
AP008209_077	7700001	7810000
AP008209_078	7800001	7910000
AP008209_079	7900001	8010000
AP008209_080	8000001	8110000
AP008209_081	8100001	8210000
AP008209_082	8200001	8310000
AP008209_083	8300001	8410000
AP008209_084	8400001	8510000
AP008209_085	8500001	8610000
AP008209_086	8600001	8710000
AP008209_087	8700001	8810000
AP008209_088	8800001	8910000
AP008209_089	8900001	9010000
AP008209_090	9000001	9110000
AP008209_091	9100001	9210000
AP008209_092	9200001	9310000
AP008209_093	9300001	9410000
AP008209_094	9400001	9510000
AP008209_095	9500001	9610000
AP008209_096	9600001	9710000
AP008209_097	9700001	9810000
AP008209_098	9800001	9910000
AP008209_099	9900001	10010000
AP008209_100	10000001	10110000
AP008209_101	10100001	10210000
AP008209_102	10200001	10310000
AP008209_103	10300001	10410000
AP008209_104	10400001	10510000
AP008209_105	10500001	10610000
AP008209_106	10600001	10710000
AP008209_107	10700001	10810000
AP008209_108	10800001	10910000
AP008209_109	10900001	11010000
AP008209_110	11000001	11110000
AP008209_111	11100001	11210000
AP008209_112	11200001	11310000
AP008209_113	11300001	11410000
AP008209_114	11400001	11510000
AP008209_115	11500001	11610000
AP008209_116	11600001	11710000
AP008209_117	11700001	11810000
AP008209_118	11800001	11910000
AP008209_119	11900001	12010000
AP008209_120	12000001	12110000
AP008209_121	12100001	12210000
AP008209_122	12200001	12310000
AP008209_123	12300001	12410000
AP008209_124	12400001	12510000
AP008209_125	12500001	12610000
AP008209_126	12600001	12710000
AP008209_127	12700001	12810000
AP008209_128	12800001	12910000
AP008209_129	12900001	13010000
AP008209_130	13000001	13110000
AP008209_131	13100001	13210000
AP008209_132	13200001	13310000
AP008209_133	13300001	13410000
AP008209_134	13400001	13510000

AP008209_135	13500001	13610000
AP008209_136	13600001	13710000
AP008209_137	13700001	13810000
AP008209_138	13800001	13910000
AP008209_139	13900001	14010000
AP008209_140	14000001	14110000
AP008209_141	14100001	14210000
AP008209_142	14200001	14310000
AP008209_143	14300001	14410000
AP008209_144	14400001	14510000
AP008209_145	14500001	14610000
AP008209_146	14600001	14710000
AP008209_147	14700001	14810000
AP008209_148	14800001	14910000
AP008209_149	14900001	15010000
AP008209_150	15000001	15110000
AP008209_151	15100001	15210000
AP008209_152	15200001	15310000
AP008209_153	15300001	15410000
AP008209_154	15400001	15510000
AP008209_155	15500001	15610000
AP008209_156	15600001	15710000
AP008209_157	15700001	15810000
AP008209_158	15800001	15910000
AP008209_159	15900001	16010000
AP008209_160	16000001	16110000
AP008209_161	16100001	16210000
AP008209_162	16200001	16310000
AP008209_163	16300001	16410000
AP008209_164	16400001	16510000
AP008209_165	16500001	16610000
AP008209_166	16600001	16710000
AP008209_167	16700001	16810000
AP008209_168	16800001	16910000
AP008209_169	16900001	17010000
AP008209_170	17000001	17110000
AP008209_171	17100001	17210000
AP008209_172	17200001	17310000
AP008209_173	17300001	17410000
AP008209_174	17400001	17510000
AP008209_175	17500001	17610000
AP008209_176	17600001	17710000
AP008209_177	17700001	17810000
AP008209_178	17800001	17910000
AP008209_179	17900001	18010000
AP008209_180	18000001	18110000
AP008209_181	18100001	18210000
AP008209_182	18200001	18310000
AP008209_183	18300001	18410000
AP008209_184	18400001	18510000
AP008209_185	18500001	18610000
AP008209_186	18600001	18710000
AP008209_187	18700001	18810000
AP008209_188	18800001	18910000
AP008209_189	18900001	19010000
AP008209_190	19000001	19110000
AP008209_191	19100001	19210000
AP008209_192	19200001	19310000
AP008209_193	19300001	19410000
AP008209_194	19400001	19510000
AP008209_195	19500001	19610000
AP008209_196	19600001	19710000

Query Match 34.4%; Score 54; DB 15; Length 110000;
Best Local Similarity 85.7%; Pred. No. 1.8e-06;
Matches 60; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 123
              ||||| |||||| ||||||| || | ||||||||||||||||||    || ||||||||||
Db      36362 CCTTGGCTATTCCGAACCAACAGAACACCTACTCCCAAACAATCACGCTGACTCATGCAA 36303

Qy        124 CTTCCATGCA 133
              | ||||||||
```

Db 36302 CCTCCATGCA 36293

RESULT 9
AC135206/c

```
LOCUS       AC135206              137327 bp    DNA     linear   PLN 16-APR-2003
DEFINITION  Oryza sativa (japonica cultivar-group) chromosome 3 clone
            OJ1041F02, complete sequence.
ACCESSION   AC135206
VERSION     AC135206.3  GI:27596977
KEYWORDS    HTG.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 137327)
  AUTHORS   Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and
            Collura,K.
  TITLE     Rice Genomic Sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 137327)
  AUTHORS   Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C.,
            Currie,J., Collura,K. and Thompson,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-OCT-2002) Arizona Genomics Institute, University of
            Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE   3  (bases 1 to 137327)
  AUTHORS   Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C.,
            Currie,J., Collura,K. and Thompson,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-NOV-2002) Arizona Genomics Institute, University of
            Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE   4  (bases 1 to 137327)
  AUTHORS   Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J.,
            Collura,K. and Thompson,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-2003) Arizona Genomics Institute, University of
            Arizona, 303 Forbes, Tucson, AZ 85721, USA
REFERENCE   5  (bases 1 to 137327)
  AUTHORS   Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T.,
            Saski,C., Henry,D., Thompson,S., Simmons,J., Thurmond,S.K. and
            Sun,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2003) Clemson University Genomics Institute,
            Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
COMMENT     On Jan 11, 2003 this sequence version replaced gi:24635891.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality
            >=30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by more than
            one plasmid subclone ; and the assembly was confirmed by
            restriction digest.There are transposons located at area
            29427-32874 and at area 19945-33212. Region 70511-70536 is covered
            by a single subclone. Region 18678-18873 is a double stranded
            single subclone.  Areas 69957-70761, 89836-90416, 98951-99009,
            105316-105607, and 133052-133152 are covered by Monsanto only.  The
            nucleotide sequence of this BAC clone was generated by combining
            Syngenta, Monsanto and Arizona Genomics Institute sequencing data.
FEATURES             Location/Qualifiers
     source          1. .137327
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39947"
                     /chromosome="3"
                     /clone="OJ1041F02"
                     /note="(japonica cultivar-group)"
     gene            complement(638. .9495)
                     /gene="OJ1041F02.1"
                     /note="unknown protein"
     CDS             complement(join(638. .838,1141. .1276,3132. .3244,3326. .3421,
                     3513. .3727,3812. .3934,4024. .4054,4157. .4216,5088. .5243,
```

```
                5322. .5411,5484. .5699,5808. .5942,6035. .6538,6622. .6837,
                7548. .7593,7679. .7815,8196. .8328,9221. .9495))
                /gene="OJ1041F02.1"
                /codon_start=1
                /product="unknown protein"
                /protein_id="AAP06844.1"
                /db_xref="GI:29893590"
                /translation="MAQSLELLLIQFLMPDNDARRQAEEQIRRLARDPQVVPALVHHL
                RTAKTPNVRQLAAVLLRKKITSHWPKLPPHAKASLKQALIDSITIDHSHLVRRASANV
                VSIIAKYAVPAGEWPELLPFIFQCSQSPQEDHREVALILFSSLTETIGTTFQSHLNDL
                QPILLKCLQDEASSRVRIAALKAVGSFIEYVNDGGDVVKIFRDFVPSILNVSRQCLAN
                GEEDVASIAFEIFDELIESPAPLLGDSVRSIVQFSLEVCSNQELEINIRQQAIQIISW
                LVKFKASFLKKHKLVIPILQVMCPLLTETADEDGDSDLAADRSAAEVIDTMAINLPRH
                VFPPVLEFASVSFRHINPKYREAAVTSLGVVSEGCCEHLKDKLEDCLKVVLEALKDQE
                QMVRGAASFALGQFAEHLQPEILSHYESVLPCILNALEDPSDEVKEKSYYALAAFCED
                MGENILPYLDPLMCRLVMSLQGSPRNLQETCMSAIGSVAAAAEQAFMPYAEKVLEMMK
                GFMVLTNDEDLCARARATEVVGIVAMAVGRARMETILPPFIEAAISGFVLDYSELREY
                THGFFSNVAEILDDSFAQYLPHVVPLAFSSCNLDDGSAVDIDDADSVDNGFSGVSSDD
                DVNDEPRVRNISWLTIIADILTAIRAIPPAHADVLEKQKDILDTVMNIYIKTMREDDD
                KEVVAQACTSLADIVRDCGFAIIEPYITRLADATLILLRQESCCQQVESDGEDDGDID
                HDEVLMDAVSDLLPAFAKVMGSYFDPIFTKLFDSLMKFAKSPHPPQDKTMVVATLAEV
                AQGMGAPISAYVDKIMPLVLKELASSEATNRRNAAFCVGEMCKNGGAAALKYYGDILH
                GLHRLFADSEPDDAVRDNAAGAIARMIMVQPQSIPLNQILPLVPDVINAFAQVVVSPN
                ESDEVKTVVAKAVSHLISVYGQQMQPILSALPPAHANALASFANRR"
gene            12372. .15503
                /gene="OJ1041F02.2"
                /note="unknown protein"
CDS             join(12372. .12757,12858. .13025,13130. .13295,13716. .13826,
                13982. .14130,14224. .14354,14459. .14558,15197. .15503)
                /gene="OJ1041F02.2"
                /codon_start=1
                /product="unknown protein"
                /protein_id="AAP06845.1"
                /db_xref="GI:29893591"
                /translation="MPPQQGECCRHGWLGVGEAAVGGGGEEPFYVPLRKRLSVDGKAS
                TAPRICIWECDGEAGDITCDIVAAPLRRSCSAKAMPPPAPLFRMMTPPPPRPQRGDGE
                EARRPGEAIRKGHRSYSLMLNLQLGISYSVGKSSALPPQKLAASDFDPREKVWTRFPP
                EGSKFTPPHHSVDFRWKDYCPAVFRHLRKLFGVDPAEYMLAICGNDTLRELASPGKSG
                SCFFITQDDRFMIKTVKKSEVKVLIRMLRSYYEHVRQYKSTLLTRFYGTHCIKQAGCP
                KIHRRFDLKGSSHGRTIDKTERKIDETTTLKDLDLQYAFRLQRFWYEELMKQIQMDCT
                FLETQGIMDYSLLLGVHFRNDYSVSKIGISQHIAFPKSTGKRKSFEGGSSFCELCFVE
                SGCKDRDLIDSRKPFIQLGINMPAQAERSSKKILDNFLLNERHLFITPPSGGSCDVYL
                FFGIIDILQDYDITKKLEHAYKSFQVNPDYISAVDPKLYSRRFQDFIRRVFIKEQ"
repeat_region   16734. .16878
                /note="putative MITE, Gaijin/Gaigin-like"
repeat_region   19263. .19454
                /note="putative transposon, CACTA, En/Spm-like"
gene            21819. .24660
                /gene="OJ1041F02.3"
                /note="hypothetical protein"
CDS             join(21819. .22007,22096. .22206,23772. .23996,24523. .24660)
                /gene="OJ1041F02.3"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAP06846.1"
                /db_xref="GI:29893592"
                /translation="MATASDGSGSGGEQRRLLSIPKEGERIIAPTRRPDGTLRKAIRI
                RAGYVPQEEVAIYQSKGAQMRKSGPDVPPGYDPALDAKPKTKAAKRNERRKEKRQQAS
                TTNDKGKGLHIEDDAGETDNPKDAVDSVTKQISGIAISESLVVATSSTDATDNSKSES
                SAPDIDKKIRALKKKIRLAEAQVQGDPENLKPEQLEKMKKIEGWKEELKLLENKSSPA
                AS"
gene            complement(26563. .30575)
                /gene="OJ1041F02.4"
                /note="unknown protein"
CDS             complement(join(26563. .26772,27082. .27288,27430. .27537,
                27881. .27940,28051. .28126,28327. .28484,28616. .28678,
                28790. .28874,29560. .30575))
                /gene="OJ1041F02.4"
                /codon_start=1
                /product="unknown protein"
                /protein_id="AAP06847.1"
                /db_xref="GI:29893593"
                /translation="MGPPAAPPSPSSSSGGSSRRRRRLERRNAAKHIGYDASNFCAYP
                QSPPAASAPASGSPSLACSPACSLDLTSFRIGGSGDGCRDVQLLCSSLGLSGVDDFAV
                PVADWEAHKAGRSSFSSSASTPKPREEPPARDSPVRREVAAEEEPPSLPAPAAAPVLP
                AKETPRSVAIEAPAPLLRVDPWEPARPDVRKASGEGGIKGVRPPPVVLKPPPSMVRPA
                VCVVESTWDILRSFAPEEDSHAHAPASRSGGDSACQDAGEEEDDAAAVLTLEELRLGE
                TSEEFTGTSSLSTTNDDETSSTTTESMFYISPNGRFRRKIRSWNRGMLLGSGSFGTVF
                EGISDEGVFFAVKEVCLCDQGSNAQQCIFQLEQEIALLSQFEHENIVQYYGTDKEDSK
                LYIFLELVTQGSLASLYQKYRLRDTHVSAYTRQILNGLTYLHERNIVHRDIKCANILV
                HANGSVKLADFGLAKEITKFNVLKSCKGTVYWMAPEVVNPKTTYGPEADIWSLGCTVL
                EMLTRQLPYPGLEWTQALYRIGKGEPPAIPNCLSRDARDFISQCVKPNPQDRPSAAKL
                LEHPFVNRSMRSIRSMRTSSRSNSSILVLLIMWSWPGRTIRYREAILSTAQSSTPTIA
                FHLLILSHSLLRGDMSWFSQRLIFSPENRLNVHSSA"
repeat_region   27679. .27832
                /note="putative MITE, MITE-adh, type D-like"
repeat_region   complement(31668. .31815)
gene            complement(35049. .49855)
                /gene="OJ1041F02.5"
                /note="hypothetical protein"
CDS             complement(join(35049. .35184,35866. .36033,36922. .37205,
                37280. .38523,40038. .40104,40196. .40287,40406. .40617,
                41678. .41760,42229. .42259,42907. .43088,43694. .43753,
                44229. .44426,44622. .44676,45877. .45992,46149. .46313,
                46447. .46620,47271. .47470,47915. .48019,48200. .48263,
                49589. .49855))
                /gene="OJ1041F02.5"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAP06848.1"
                /db_xref="GI:29893594"
                /translation="MGGSADPEAPTPTPSPSPSPAKATPSPASADGNRLRRCVQSKLS
                WGPPKAGGGGGEAGGAGLPPLAAGDGTPEKVKKRGRPRKSEAGKKPSSNRETTGLEQD
                SKDEVILVDESPQKKQRKGRGKNQGAALKVPNRKHCKALESTDGHESCQQLRSSQTQA
                VLPQKSPTSVDIDLVTGPSEASPVNDNVDALDNEDKPQLIVDLRSEANIAAEENRRLS
                SGKKMHPFFASRKIHKGAGQDILNVEDEDMDSLCAFERDPPLCPVHVLYELEVTMPIH
                WSNKWLIADKSFLGTSTTEQNSAEHADPGKHLANFHDKQNKSKFSSQDVIDVDDECLL
                ASSSCFHASLFESKQHERVQHELPEVTPKGCQTANLWTDKYRPETAAQVCGNSKHVKF
                LNEWLKGWDERGHRNKQNIVTNGSMNGRSCQDGSDTDYSEDASDYENVLLITGPVGCG
                KSAAVFACAREQGFNVIELNMPKTVVFPFLPHNKLLAQTLDLKPDKLHDSSHPSIKYI
                PTRVFVHMNRNSIISISAHSSSLSVIKLHCLVNTSDMRNGAYVRQKFEEATKSHGLEK
                WSQEEIIGLPISNSLDPASGTPGTAEYKQVINKTLILFEDVDTVFDEDRGFISTILKM
                VETTKWPIILTSNKKDPPLPHLLAQLVLDFTYPSSAELLSHVDMICKSEGVEITVPQQ
                KHIIDAFLGRLNKCLSCPSLLDLDAVHSTVPRIMPWDFPCKLSETIYMEIDKTIVTAE
                QKKKQMEVSEFEGLELQIMTPLTKGRSAGKTRKPKKSKLKHGRSADCNDASPCKNDLD
                DFHDSPDIPLPSNHQRMRNRRGVVLFAESDDDLADAHAAKDATFTVQEGRLLPQSSEL
                PCLYGHGISNIVPESVFFQQSSVPHLHREVISNQLCFPSESRAFEPASSFQNQLESNM
                PGSISQICDTFMSQGISCVPESSFMVGGTSASISSDDLLSSLVSNGLSALRNESTYTA
                SVVALEDTNKVENQMTDKPQKCMEDEVGETCEAYVELADRNDHASCSITGYQLMDECS
                RAESVWLLSGKKNNDSCKVEHVQDTWNRLRQCHPVLPCDMNHNRSVSGALKRVSRVSD
                LISESDLMLISCHPFSNDISDPSLTPYTESDGFSYSKQLEMGSIYAQHGLCIFLQDSQ
                ATDDGFVDFLQELLFSGTTTTSLGKFVSSGISCGDGSGNISHVKYPTSCISKRRERQA
                RLREVLLPVVPPKLSQSLRGPAFVDYLSSMSQISQLENMQLSECKASSKQRRCRQPRH
                YLSSGALSLSAEDIGLLAQCSTFSDRRESETIIEQAIS"
repeat_region   39012. .39207
                /note="putative MITE, MITE-adh, type M-like"
repeat_region   39681. .39879
                /note="putative MITE, MITE-adh, type G-like"
repeat_region   51474. .51836
                /note="putative MITE, Castaway-like"
gene            complement(59793. .62196)
                /gene="OJ1041F02.6"

  Query Match             34.4%;  Score 54;  DB 15;  Length 137327;
  Best Local Similarity   85.7%;  Pred. No. 1.8e-06;
  Matches   60;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 123
              ||||| |||||| ||||||| || | ||||||||||||||||||   || ||||||||||
Db     119499 CCTTGGCTATTCCGAACCAACAGAACACCTACTCCCAAACAATCACGCTGACTCATGCAA 119440

Qy        124 CTTCCATGCA 133
              | ||||||||
Db     119439 CCTCCATGCA 119430
```

```
RESULT 10
AX224398
LOCUS       AX224398                  50 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 5 from Patent WO0160997.
ACCESSION   AX224398
VERSION     AX224398.1  GI:15554640
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 5 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match           31.8%;  Score 50;  DB 6;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-05;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         75 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 50


RESULT 11
AX224399
LOCUS       AX224399                  40 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 6 from Patent WO0160997.
ACCESSION   AX224399
VERSION     AX224399.1  GI:15554641
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 6 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match           25.5%;  Score 40;  DB 6;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 0.066;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         85 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||||||||||||||||||||||||||||||||||||||
Db          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40


RESULT 12
BX950854/c
LOCUS       BX950854              171896 bp    DNA     linear   VRT 03-FEB-2005
DEFINITION  Zebrafish DNA sequence from clone CH211-125M22, complete sequence.
ACCESSION   BX950854
VERSION     BX950854.12  GI:58190604
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 171896)
  AUTHORS   Pelan,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 25, 2005 this sequence version replaced gi:56368832.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            CH211-125M22 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .171896
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-125M22"
                     /clone_lib="CHORI-211"
ORIGIN

  Query Match           22.2%;  Score 34.8;  DB 5;  Length 171896;
  Best Local Similarity   52.8%;  Pred. No. 3.1;
  Matches   75;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy         15 TGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATT 74
              |||  |  | |||  ||||  | | ||  | |   ||||  | |  ||  ||||||    | |
Db     158604 TGGGTTTGTCCATACTTGACCCAACGCTGGGTGTGTTAAAACAACCCAGCCTTTTAAAGT 158545

Qy         75 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAA 134
                | | ||      ||   |  || ||| | | ||| | |    || ||  ||  |  ||
Db     158544 AAGTAACATTCATCTAAAAATGCCTAAATATTTCATTTAAAATCTGAAATCTCAGTTTAA 158485

Qy        135 ACACGCACATATGTTTCCTGAA 156
              || | ||||| | |    | ||
Db     158484 ACTCCCACATTTCTCATTTAAA 158463
```

```
RESULT 13
AC104134
LOCUS       AC104134              110127 bp    DNA     linear   PRI 30-APR-2005
DEFINITION  Homo sapiens BAC clone RP11-525L16 from 2, complete sequence.
ACCESSION   AC104134
VERSION     AC104134.4  GI:19551199
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 110127)
  AUTHORS   Doebber,A., Haakenson,W. and Tomlinson,C.
  TITLE     The sequence of Homo sapiens BAC clone RP11-525L16
  JOURNAL   Unpublished (2001)
REFERENCE   2  (bases 1 to 110127)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   3  (bases 1 to 110127)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 110127)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 110127)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-APR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Mar 20, 2002 this sequence version replaced gi:18677630.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0525L16
            --------------.

            NOTICE:

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. Wes Warren,
            Department of Genetics, Washington University, St. Louis MO. For
            additional information about the map position of this sequence, see
            http://genome.wustl.edu

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-554H10, 2000 bp overlap;
            the clone sequenced to the right is RP11-450E9, 2000 bp overlap.
            Actual start of this clone is at base position 151996 of
            RP11-554H10.
FEATURES              Location/Qualifiers
     source           1. .110127
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="2"
                      /clone="RP11-525L16"
                      /clone_lib="RPCI-11"
     misc_feature     4046. .4309
                      /note="CpG island (%GC=61.0, o/e=0.83, #CpGs=20)"
     misc_feature     13008. .13484
                      /note="CpG island (%GC=75.1, o/e=0.69, #CpGs=53)"
     gene             34638. .39571
                      /gene="FLJ25369"
     mRNA             join(34638. .34769,35622. .35658,36392. .36484,39039. .39571)
                      /gene="FLJ25369"
     CDS              join(35632. .35658,36392. .36484,39039. .39461)
                      /gene="FLJ25369"
                      /note="Homo sapiens hypothetical protein FLJ25369
                      (FLJ25369), mRNA.; H_NH0525L16.1
                      This gene was based on gi(22749356)"
                      /codon_start=1
                      /product="unknown"
                      /protein_id="AAY24330.1"
                      /db_xref="GI:62988943"
                      /translation="MAGVKYPGQDPVDLDIYQSSHMVDYQPYRKHKYSRVTPQEQAKL
                      DAQLRDKEFYRPIPNPNPKLTDGYPAFKRPHMTAKDLGLPGFFPSQEHEATREDERKF
                      TSTCHFTYPASHDLHLAQGDPNQVLQSADFPCLVDPKHQPAAEMAKGYLLLPGCPCLH
                      CHIVKVPILNRWGPLMPFYQ"
     misc_feature     44296. .44562
                      /note="CpG island (%GC=62.5, o/e=0.89, #CpGs=24)"
     misc_feature     59905. .60161
                      /note="CpG island (%GC=66.5, o/e=0.76, #CpGs=22)"
     gene             complement(66728. .105650)
                      /gene="EIF2AK3"
     mRNA             complement(join(66728. .67923,68904. .68966,72276. .72377,
                      80861. .81028,84653. .85433,86541. .86690,89505. .89627,
                      93417. .93529,95828. .96048,97969. .98091,98748. .98888,
                      100430. .100592,100805. .101039,103259. .103392,
                      105456. .105650))
                      /gene="EIF2AK3"
     CDS              complement(join(67723. .67923,68904. .68966,72276. .72377,
                      80861. .81028,84653. .85433,86541. .86690,89505. .89627,
                      93417. .93529,95828. .96048,97969. .98091,98748. .98888,
                      100430. .100592,100805. .101039,103259. .103392,
                      105456. .>105650))
                      /gene="EIF2AK3"
                      /note="Homo sapiens eukaryotic translation initiation
                      factor 2-alpha kinase 3 (EIF2AK3), mRNA.; H_NH0525L16.2
                      This gene was based on gi(21361154)
                      Continued from H_NH0450E09.1"
                      /codon_start=1
                      /product="unknown"
                      /protein_id="AAY24331.1"
                      /db_xref="GI:62988944"
                      /translation="VFGNKMIIPSLDGALFQWDQDRESMETVPFTVESLLESSYKFGD
                      DVVLVGGKSLTTYGLSAYSGKVRYICSALGCRQWDSDEMEQEEDILLLQRTQKTVRAV
                      GPRSGNEKWNFSVGHFELRYIPDMETRAGFIESTFKPNENTEESKIISDVEEQEAAIM
                      DIVIKVSVADWKVMAFSKKGGHLEWEYQFCTPIASAWLLKDGKVIPISLFDDTSYTSN
                      DDVLEDEEDIVEAARGATENSVYLGMYRGQLYLQSSVRISEKFPSSPKALESVTNENA
                      IIPLPTIKWKPLIHSPSRTPVLVGSDEFDKCLSNDKFSHEEYSNGALSILQYPYDNGY
                      YLPYYKRERNKRSTQITVRFLDNPHYNKNIRKKDPVLLLHWWKEIVATILFCIIATTF
                      IVRRLFHPHPHRQRKESETQCQTENKYDSVSGEANDSSWNDIKNSGYISRYLTDFEPI
                      QCLGRGGFGVVFEAKNKVDDCNYAIKRIRLPNRELAREKVMREVKALAKLEHPGIVRY
                      FNAWLEAPPEKWQEKMDEIWLKDESTDWPLSSPSPMDAPSVKIRRMDPFATKEHIEII
```

APSPQRSRSFSVGISCDQTSSSESQFSPLEFSGMDHEDISESVDAAYNLQDSCLTDCD
VEDGTMDGNDEGHSFELCPSEASPYVRSRERTSSSIVFEDSGCDNASSKEEPKTNRLH
IGNHCANKLTAFKPTSSKSSSEATLSISPPRPTTLSLDLTKNTTEKLQPSSPKVYLYI
QMQLCRKENLKDWMNGRCTIEERERSVCLHIFLQIAEAVEFLHSKGLMHRDLKPSNIF
FTMDDVVKVGDFGLVTAMDQDEEEQTVLTPMPAYARHTGQVGTKLYMSPEQIHGNSYS
HKVDIFSLGLILFELLYPFSTQMERVRTLTDVRNLKFPPLFTQKYPCEYVMVQDMLSP
SPMERPEAINIIENAVFEDLDFPGKTVLRQRSRSLSSSGTKHSRQSNNSHSPLPSN"

ORIGIN

Query Match 22.0%; Score 34.6; DB 8; Length 110127;
Best Local Similarity 54.3%; Pred. No. 3.6;
Matches 70; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

```
Qy          3 TGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAA 62
              || ||  ||   |  |  |    | | ||  |||||  || | || ||||    | | |
Db      87633 TGGGATTACAAGCGTGAGCCACTGTACCTGGCCAACCTATCTTTTTCTTCCTCTCTTCCA 87692

Qy         63 GCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCA 122
                |||  |||||||| | ||        | |||||||  |||. ||    ||   ||| |
Db      87693 TCCTCCCCTATTCTTACCCTTTGTTCCATCTACTCCAGAACTCTCACCATTCTCCATACC 87752

Qy        123 ACTTCCATG 131
                || |||| |
Db      87753 ACCTCCAAG 87761
```

RESULT 14
AC136896
LOCUS AC136896 214946 bp DNA linear PRI 21-MAR-2003
DEFINITION Homo sapiens chromosome 15, clone RP11-86001, complete sequence.
ACCESSION AC136896
VERSION AC136896.6 GI:29135656
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 214946)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 15, clone RP11-86001
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 214946)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (09-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 214946)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (23-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 214946)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Mar 21, 2003 this sequence version replaced gi:27877278.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
---------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
---------- Project Information
Center project name: L27824
Center clone name: 860_O_1

FEATURES Location/Qualifiers
source 1. .214946
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-86001"
/clone_lib="RPCI-11 Human Male BAC"
repeat_region complement(1. .362)
/rpt_family="HAL1"
repeat_region complement(823. .1431)
/rpt_family="L1MC5"
unsure 913. .922
/note="<30 qual SNGL region"
repeat_region complement(2941. .3295)
/rpt_family="MER31B"
unsure complement(3307. .3311)
/note="<30 qual SNGL region"
unsure 3438. .3476
/note="single clone coverage"

```
repeat_region   3546. .3586
                /rpt_family="GA-rich"
repeat_region   3601. .4006
                /rpt_family="MLT1A1"
repeat_region   complement(4008. .4166)
                /rpt_family="MIR"
repeat_region   complement(4580. .4711)
                /rpt_family="L3"
repeat_region   4731. .4751
                /rpt_family="AT_rich"
repeat_region   complement(4758. .4993)
                /rpt_family="MIR"
repeat_region   complement(5336. .5464)
                /rpt_family="L3"
repeat_region   5935. .6412
                /rpt_family="MER39"
repeat_region   complement(6466. .7069)
                /rpt_family="MLT1E"
repeat_region   complement(8633. .8713)
                /rpt_family="MIR"
repeat_region   8768. .9244
                /rpt_family="L1MA2"
repeat_region   complement(9299. .9498)
                /rpt_family="MIR"
repeat_region   10252. .10587
                /rpt_family="MLT1A0"
repeat_region   12232. .12619
                /rpt_family="L1MA5"
repeat_region   complement(13183. .13354)
                /rpt_family="MIR"
repeat_region   14074. .14500
                /rpt_family="L1MB3"
repeat_region   15070. .15201
                /rpt_family="FLAM_C"
repeat_region   complement(16057. .18621)
                /rpt_family="L1MB3"
repeat_region   complement(18678. .19377)
                /rpt_family="L1M1"
repeat_region   complement(19375. .20142)
                /rpt_family="L1MB3"
repeat_region   22615. .22919
                /rpt_family="AluSg"
repeat_region   complement(25101. .25206)
                /rpt_family="HAL1"
repeat_region   complement(25691. .25925)
                /rpt_family="HAL1"
repeat_region   25996. .26033
                /rpt_family="(CAGA)n"
repeat_region   complement(26059. .26582)
                /rpt_family="MLT2B3"
repeat_region   26645. .27455
                /rpt_family="MER21B"
repeat_region   complement(27582. .27833)
                /rpt_family="HAL1"
repeat_region   complement(28185. .28759)
                /rpt_family="LTR68"
repeat_region   complement(29229. .29603)
                /rpt_family="THE1A"
repeat_region   complement(29771. .30036)
                /rpt_family="AluJb"
repeat_region   30056. .30083
                /rpt_family="(TTTA)n"
repeat_region   complement(30090. .30369)
                /rpt_family="AluJo"
repeat_region   31042. .32585
                /rpt_family="L1M4"
repeat_region   32586. .32654
                /rpt_family="(TCTA)n"
repeat_region   32655. .33163
                /rpt_family="L1M4"
repeat_region   33164. .33473
                /rpt_family="AluY"
repeat_region   33474. .36683
                /rpt_family="L1M4"
repeat_region   36696. .36751
                /rpt_family="MLT1D"
repeat_region   36941. .37015
                /rpt_family="MIR"
repeat_region   complement(38001. .38449)
                /rpt_family="L1MD1"
repeat_region   38453. .38855
                /rpt_family="MLT1D"

  Query Match           21.9%;  Score 34.4;  DB 8;  Length 214946;
  Best Local Similarity   54.8%;  Pred. No. 4.2;
  Matches   68;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy          5 TCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGC 64
              ||||||| |||| | | || || ||| | || || |||||| | || |
Db      35745 TCATCTCTATGGGCAGAAAAAATCATTTGACAAAATTACGCATTCTTCCATGATGTAAAC 35804

Qy         65 CTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
               |  || || |  |||| | || ||     |  |||| || | ||  ||| ||
Db      35805 ATCCAACATACTAAGAAAAGAAGGTAGCTTTCTTAACCCAATAAATGTAACCTATGAAAA 35864

Qy        125 TTCC 128
                |||
Db      35865 ATCC 35868


RESULT 15
CR847532
LOCUS       CR847532              141554 bp    DNA     linear   HTG 15-JAN-2005
DEFINITION  Danio rerio clone DKEYP-86C9, WORKING DRAFT SEQUENCE, 3 unordered
            pieces.
ACCESSION   CR847532
VERSION     CR847532.4  GI:57863691
KEYWORDS    HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 141554)
  AUTHORS   Pelan,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 15, 2005 this sequence version replaced gi:56309949.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zKp86C9
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 141058 bases at least Q40
            Consensus quality: 141224 bases at least Q30
            Consensus quality: 141289 bases at least Q20
            Insert size: 141354; sum-of-contigs
            Insert size: 152619; 10.9% error; agarose-fp
            Quality coverage: 8.68x in Q20 bases; sum-of-contigs Quality
            coverage: 8.12x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
```

```
            *        1    10378: contig of 10378 bp in length
            *    10379    10478: gap of 100 bp
            *    10479    42327: contig of 31849 bp in length
            *    42328    42427: gap of 100 bp
            *    42428   141554: contig of 99127 bp in length.
FEATURES             Location/Qualifiers
     source          1. .141554
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEYP-86C9"
                     /clone_lib="DanioKeypilot"
     misc_feature    1. .10378
                     /note="assembly_fragment:00017"
     misc_feature    10479. .42327
                     /note="assembly_fragment:00290"
     misc_feature    42428. .141554
                     /note="assembly_fragment:00621.0"
ORIGIN

  Query Match             21.5%;  Score 33.8;  DB 14;  Length 141554;
  Best Local Similarity   52.5%;  Pred. No. 6.6;
  Matches   74;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy         12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
              ||| |    |||||| ||||| | |  | | | |||| | |  || ||    ||| |
Db      90222 ACAGTGACCCCTACATCCTTGTGCCTCAATGCATATTGTACAAGTGTACAGAGGTTGGGT 90281

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              |   | |  | | ||||   | ||||  || |||  |  | | |||  |||       |
Db      90282 AAGTGGGAATATGGGGATCATTGCTCCTGAAAAATGAAAATAAATCACACAAAGAAACTT 90341

Qy        132 CAAACACGCACATATGTTTCC 152
              || |   || |||   |||
Db      90342 TCAAAAAAAACTTATTAGTCC 90362


Search completed: March  5, 2006, 21:55:32
Job time : 690.29 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:08:23 ; Search time 104.48 Seconds
(without alignments)
10014.946 Million cell updates/sec

Title: US-10-713-381-1_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	157	100.0	255	5	AAH76340	Aah76340 Z. mays M
	2	157	100.0	1394	2	AAX07408	Aax07408 Zea mays
	3	157	100.0	1394	2	AAX07409	Aax07409 Zea mays
	4	157	100.0	1394	5	AAH76332	Aah76332 Z. mays M
	5	157	100.0	1394	5	AAH76333	Aah76333 Z. mays M
	6	146	93.0	158	5	AAH76334	Aah76334 Z. mays M
	7	54	34.4	2000	11	ACL38730	Acl38730 Rice stre
	8	50	31.8	50	5	AAH76336	Aah76336 Z. mays M
	9	40	25.5	40	5	AAH76337	Aah76337 Z. mays M
	10	32.4	20.6	3267	12	ADK52131	Adk52131 Mouse ato
c	11	31.6	20.1	2352	8	ACA48402	Aca48402 Prokaryot
c	12	31.4	20.0	15734	6	ABN80329	Abn80329 Human che
	13	30.6	19.5	64415	3	AAF22279	Aaf22279 BAC conta
	14	30.6	19.5	79467	9	ADA02717	Ada02717 Mouse Nfa
	15	30.6	19.5	79467	10	ADB72455	Adb72455 Mouse Nfa
	16	30.6	19.5	79467	10	ADE95965	Ade95965 Mouse Nfa
	17	30.4	19.4	52001	12	ADP67076	Adp67076 Human mit
c	18	30.2	19.2	669	3	AAA51436	Aaa51436 A. thalia
	19	30.2	19.2	1266	9	ADB08983	Adb08983 Alloiococ
	20	30.2	19.2	110000	9	ADB12064_07	Continuation (8 of
	21	30	19.1	30	5	AAH76335	Aah76335 Z. mays M
c	22	30	19.1	378	10	ADD33403	Add33403 Mouse mit
	23	30	19.1	396	10	ADD33404	Add33404 Mouse mit
c	24	29.8	19.0	373	4	AAH13328	Aah13328 Human cDN
c	25	29.6	18.9	1185	10	ADF00796	Adf00796 Bacterial
c	26	29.4	18.7	110000	3	AAF22303_2	Continuation (3 of
c	27	29.2	18.6	1475	9	AAL62902	Aal62902 Vernonia
c	28	29.2	18.6	110000	2	AAT42063_06	Continuation (7 of
c	29	28.8	18.3	1500	8	ADA68284	Ada68284 Arabidops
	30	28.6	18.2	756	13	ADS57339	Ads57339 Bacterial
c	31	28.6	18.2	778	4	AAS32892	Aas32892 Human gen
c	32	28.6	18.2	778	4	AAS32893	Aas32893 Human gen
	33	28.6	18.2	1043	10	ADD71973	Add71973 Human uri
	34	28.6	18.2	1522	13	ADW26566	Adw26566 Glycine m
	35	28.6	18.2	1860	6	ABN67847	Abn67847 Streptoco
c	36	28.6	18.2	2160	13	ADX11599	Adx11599 Plant ful
c	37	28.6	18.2	3231	13	ADX13304	Adx13304 Plant ful
c	38	28.6	18.2	51615	11	ACN45162	Acn45162 Human gen
	39	28.6	18.2	110000	13	ABD32806_5	Continuation (6 of
c	40	28.6	18.2	164702	8	ACF62730	Acf62730 Cancer ba
c	41	28.6	18.2	164702	8	ADB20845	Adb20845 MRP1 base
c	42	28.6	18.2	164702	10	ADB87934	Adb87934 Human UGT
c	43	28.6	18.2	164702	10	ADB96917	Adb96917 Human MDR
c	44	28.6	18.2	164702	10	ADB92108	Adb92108 Human MDR
	45	28.4	18.1	1478	3	AAF15843	Aaf15843 Human pro

ALIGNMENTS

RESULT 1
AAH76340
ID AAH76340 standard; DNA; 255 BP.
XX
AC AAH76340;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 promoter fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; promoter; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Example 5; Fig 8; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The

```
CC   present sequence represents a Z. mays Ms45 promoter fragment
XX
SQ   Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 157;  DB 5;  Length 255;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-42;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 74

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 134

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db        135 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171


RESULT 2
AAX07408
ID   AAX07408 standard; DNA; 1394 BP.
XX
AC   AAX07408;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW   Ms45; male; tissue-preferred; regulatory region; plant cells;
KW   plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS   Zea mays.
XX
PN   WO9859061-A1.
XX
PD   30-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US012895.
XX
PR   23-JUN-1997;   97US-00880499.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman GA,  Kendall TL;
XX
DR   WPI; 1999-105628/09.
XX
PT   New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT   - useful in mediating plant fertility, especially hybrid seed production.
XX
PS   Claim 2; Page 22-23; 39pp; English.
XX
CC   The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC   region. It may be used in the construction of a vector for a method of
CC   producing exogenous genes in a male tissue- preferred manner, which is
CC   useful in restoring or conferring fertility, such as in hybrid seed
CC   production. In conferring fertility, a monocot/dicot plant is transformed
CC   with the exogenous nucleotide sequence (a male sterility gene, preferably
CC   Ms45), which encodes a product selected from auxins, rolB and diptheria
CC   toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 157;  DB 2;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-42;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 3
AAX07409
ID   AAX07409 standard; DNA; 1394 BP.
XX
AC   AAX07409;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Zea mays Ms45 male tissue-preferred regulatory region.
XX
KW   Ms45; male; tissue-preferred; regulatory region; plant cells;
KW   plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS   Zea mays.
XX
PN   WO9859061-A1.
XX
PD   30-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US012895.
XX
PR   23-JUN-1997;   97US-00880499.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman GA,  Kendall TL;
XX
DR   WPI; 1999-105628/09.
XX
PT   New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT   - useful in mediating plant fertility, especially hybrid seed production.
XX
PS   Claim 3; Page 23-24; 39pp; English.
XX
CC   The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC   region. It may be used in the construction of a vector for a method of
CC   producing exogenous genes in a male tissue- preferred manner, which is
CC   useful in restoring or conferring fertility, such as in hybrid seed
CC   production. In conferring fertility, a monocot/dicot plant is transformed
CC   with the exogenous nucleotide sequence (a male sterility gene, preferably
CC   Ms45), which encodes a product selected from auxins, rolB and diptheria
CC   toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC   and infertile plants
XX
SQ   Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 157;  DB 2;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-42;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

```
RESULT 4
AAH76332
ID   AAH76332 standard; DNA; 1394 BP.
XX
AC   AAH76332;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 4; Page 46; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a nucleic acid sequence encoding an Ms45 male
CC   -tissue preferred regulatory region from Z. mays
XX
SQ   Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 157;  DB 5;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-42;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 5
AAH76333
ID   AAH76333 standard; DNA; 1394 BP.
XX
AC   AAH76333;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 4; Page 47; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a nucleic acid sequence encoding an Ms45 male
CC   -tissue preferred regulatory region from Z. mays
XX
SQ   Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 157;  DB 5;  Length 1394;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-42;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 6
AAH76334
ID   AAH76334 standard; DNA; 158 BP.
XX
AC   AAH76334;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
```

```
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
PT   A male tissue-preferred regulatory region comprising nucleotide sequences
PT   essential for initiating transcription of the MS45 gene useful for
PT   mediating fertility in a male plant.
XX
PS   Claim 5; Page 47; 50pp; English.
XX
CC   The invention provides a male tissue-preferred regulatory region (I)
CC   comprising nucleotide sequences essential for initiating transcription of
CC   the MS45 gene. A method of mediating male fertility in a plant is
CC   provided that involves introducing an expression vector comprising a
CC   promoter operably linked to (I) into a plant where the exogenous gene
CC   impacts male fertility of the plant and (I) controls expression of the
CC   exogenous gene. A method of producing hybrid seeds is also provided. The
CC   present sequence represents a DNA fragment -38 to -195 bases upstream of
CC   the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC   nucleotide sequence
XX
SQ   Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

  Query Match             93.0%;  Score 146;  DB 5;  Length 158;
  Best Local Similarity   99.4%;  Pred. No. 6.1e-39;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 59
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy         60 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120

Qy        120 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158


RESULT 7
ACL38730
ID   ACL38730 standard; cDNA; 2000 BP.
XX
AC   ACL38730;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Rice stress-regulated promoter SEQ ID NO:17293.
XX
KW   ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW   agriculture.
XX
OS   Oryza sativa.
XX
PN   WO2003008540-A2.
XX
PD   30-JAN-2003.
XX
PF   21-JUN-2002; 2002WO-US019668.
XX
PR   22-JUN-2001; 2001US-0300112P.
PR   24-AUG-2001; 2001US-0314662P.
PR   26-SEP-2001; 2001US-0325277P.
PR   21-NOV-2001; 2001US-0332132P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Kreps J,  Briggs SP,  Cooper B,  Glazebrook J,  Goff SA,  Katagiri F;
PI   Moughamer T,  Provart N,  Ricke D,  Zhu T;
XX
DR   WPI; 2003-248011/24.
XX
PT   New stress-responsive nucleic acid, useful for altering the
PT   responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT   stress, salt stress or osmotic stress.
XX
PS   Claim 48; SEQ ID NO 17293; 89pp; English.
XX
CC   The invention relates to novel abiotic stress responsive polynucleotides
CC   and polypeptides. Also disclosed are vectors, expression cassettes, host
CC   cells, and plants containing such polynucleotides. Also disclosed are
CC   methods for using the polynucleotides and polypeptides to alter the
CC   responsiveness of a plant to abiotic stress. The invention is useful in
CC   agriculture. The nucleic acid is useful for determining whether a test
CC   plant has been exposed to an abiotic stress condition. It is also useful
CC   for selecting an agent that alters abiotic stress regulated
CC   polynucleotide expression in a plant cell, and to identify a homolog or
CC   ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC   molecule and the polypeptide encoded by it are useful in altering the
CC   responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC   stress, osmotic stress or any of their combinations. The present sequence
CC   is used in the exemplification of the invention
XX
SQ   Sequence 2000 BP; 616 A; 370 C; 355 G; 659 T; 0 U; 0 Other;

  Query Match             34.4%;  Score 54;  DB 11;  Length 2000;
  Best Local Similarity   85.7%;  Pred. No. 1.2e-07;
  Matches   60;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 123
              ||||| |||||| ||||||| || | ||||||||||||||||||    || ||||||||||
Db       1819 CCTTGGCTATTCCGAACCAACAGAACACCTACTCCCAAACAATCACGCTGACTCATGCAA 1878

Qy        124 CTTCCATGCA 133
              | ||||||||
Db       1879 CCTCCATGCA 1888


RESULT 8
AAH76336
ID   AAH76336 standard; DNA; 50 BP.
XX
AC   AAH76336;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW   Ms45; male tissue; regulatory region; transcription; male fertility;
KW   hybrid seed; ds.
XX
OS   Zea mays.
XX
PN   WO200160997-A2.
XX
PD   23-AUG-2001.
XX
PF   13-FEB-2001; 2001WO-US004527.
XX
PR   15-FEB-2000; 2000US-00504487.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Albertsen MC,  Fox TW,  Garnaat CW,  Huffman G,  Kendall TL;
XX
DR   WPI; 2001-514772/56.
XX
```

PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -72 to -111 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;

Query Match 31.8%; Score 50; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 50

RESULT 9
AAH76337
ID AAH76337 standard; DNA; 40 BP.
XX
AC AAH76337;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 14; Page 32; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment upstream of the TATA box of a
CC Z. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX
SQ Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;

Query Match 25.5%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
||||||||||||||||||||||||||||||||||||||||
Db 1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40

RESULT 10
ADK52131
ID ADK52131 standard; cDNA; 3267 BP.
XX
AC ADK52131;
XX
DT 20-MAY-2004 (first entry)
XX
DE Mouse atopic-dermatitis/psoriasis-associated EST #4.
XX
KW Mouse; ss; EST; atopic dermatitis; psoriasis; dermatological;
KW anti-inflammatory; antipsoriatic; rash; expressed sequence tag.
XX
OS Mus musculus.
XX
PN WO2004016785-A1.
XX
PD 26-FEB-2004.
XX
PF 06-AUG-2003; 2003WO-JP009999.
XX
PR 06-AUG-2002; 2002JP-00229319.
PR 14-MAY-2003; 2003JP-00136544.
XX
PA (GENO-) GENOX RES INC.
PA (UYJU-) UNIV JUNTENDO.
XX
PI Itoh M, Ogawa K, Shinagawa A, Sudo H, Ogawa H, Ra C;
PI Mitsuishi K;
XX
DR WPI; 2004-214514/20.
XX
PT Detecting atopic dermatitis or psoriasis comprises assaying levels of
PT expression of an indicator gene at a rash site and non-rash site of a
PT person with atopic dermatitis or psoriasis.
XX
PS Claim 20; SEQ ID NO 164; 484pp; Japanese.
XX
CC The invention relates to detecting atopic dermatitis or psoriasis
CC comprising assaying the levels of expression of an indicator gene at a
CC rash site and non-rash site of a person with atopic dermatitis or
CC psoriasis, comparing these levels with those of a healthy person, and
CC determining that if the levels of indicators are higher or lower, then
CC this indicates the disease. Also included are a reagent for detecting
CC atopic dermatitis or psoriasis, a kit for screening for treatments, a
CC transgenic non human vertebrate animal models for the diseases, an agent
CC for inducing the diseases in mice and a DNA chip for assaying for the
CC indicator genes. The method is used for treatment, detection and animal
CC models for research of atopic dermatitis and psoriasis. The present
CC sequence is a Mouse atopic-dermatitis/ psoriasis-associated EST
CC (expressed sequence tag).
XX
SQ Sequence 3267 BP; 1029 A; 658 C; 590 G; 990 T; 0 U; 0 Other;

Query Match 20.6%; Score 32.4; DB 12; Length 3267;
Best Local Similarity 54.1%; Pred. No. 2.5;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 27 TGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAG 86
Db 2774 TTCTTTTTCTCACCTCCCTCTGATTTCTTTGTACCTCACTTTCATTTTGGCCTTTAAGTT 2833

```
Qy         87 GATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATAT 146
              | |    | ||||||| || |||| || | |       |   ||| | ||||| |  | ||
Db       2834 GTTCTAGATTCCCAAAGAAGCCATTTTGCACTACAGTCAAACATACTAACACCCTGAAAT 2893

Qy        147 GT 148
              ||
Db       2894 GT 2895
```

RESULT 11
ACA48402/c
ID ACA48402 standard; DNA; 2352 BP.
XX
AC ACA48402;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #30059.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Streptococcus mutans.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU44532.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 36272; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2352 BP; 762 A; 401 C; 500 G; 689 T; 0 U; 0 Other;

```
  Query Match            20.1%;  Score 31.6;  DB 8;  Length 2352;
  Best Local Similarity  58.5%;  Pred. No. 4.2;
  Matches   55;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy         62 AGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGC 121
              | |||  |      |  || | | ||||  | ||||||  || ||||  |  |||   |||
Db        234 ATCCTGACTGTGCCCAAAGCCAAAGGAAGCATACTCCTGAAAAATCTGTGCTACATCTGC 175

Qy        122 AACTTCCATGCAAACACGCACATATGTTTCCTGA 155
              || |||     ||  |||  ||  | ||||||||
Db        174 AATTTCAGCAAAAGAACGTTCAATTTTTTCCTGA 141
```

RESULT 12
ABN80329/c
ID ABN80329 standard; DNA; 15734 BP.
XX
AC ABN80329;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human chemically modified disease associated gene SEQ ID NO 346.
XX
KW Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW antidiabetic; cytostatic; anticonvulsant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200927-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007536.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130908/17.
XX
PT Novel nucleic acid useful for diagnosis and therapy of diseases
PT associated with development genes such as diabetes, comprises a sequence
PT of a segment of chemically pretreated DNA of genes associated with
PT development.
XX
PS Claim 1; SEQ ID NO 346; 27pp; English.
XX
CC The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases in length of a segment of chemically pretreated DNA (II)
CC of genes associated with development selected from 87 genes listed in the
CC specification such as ACCPN, ADFN, or AFD1 and comprising one of 350
CC sequences (ABN79984-ABN80333) or their complements. The invention is
CC useful for the diagnosis or therapy of diseases associated with
CC development genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to

```
CC   histone deacetylation, Currarino syndrome, diseases related with the
CC   development of the brain and limb girdle muscular dystrophy and dwarfism.
CC   Oligomers specific to each of the genes are useful for detecting the
CC   methylation state of all CpG dinucleotides within the 350 sequences or
CC   (II) and their complementary sequences, as primer oligonucleotides for
CC   the amplification of the 350 sequences, (II) and/or their complements and
CC   as oligomer probes for detecting the cytosine methylation state and/or
CC   single nucleotide polymorphisms (SNPs). Note: The sequence data for this
CC   patent did not form part of the printed specification but is based on
CC   sequence information supplied to Derwent by the European Patent Office
XX
SQ   Sequence 15734 BP; 3630 A; 744 C; 4234 G; 7126 T; 0 U; 0 Other;

  Query Match            20.0%;  Score 31.4;  DB 6;  Length 15734;
  Best Local Similarity  59.6%;  Pred. No. 9;
  Matches   53;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy         52 CCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATC 111
              || || | |   ||||  | | ||| || |||    |  ||||     ||||||| | | |
Db       8802 CCTTCTTTCTCCCCTTATCGAATCTTAAACAATCAAAACCCTAAAATAAAACAACCAAAC 8743

Qy        112 TTACTCATGCAACTTCCATGCAAACACGC 140
              |  || |  || ||||||   |||||| |
Db       8742 TACCTAACCCATCTTCCAAATAAACACCC 8714


RESULT 13
AAF22279
ID   AAF22279 standard; DNA; 64415 BP.
XX
AC   AAF22279;
XX
DT   20-MAR-2001  (first entry)
XX
DE   BAC containing repeats from centromeres 1-4 #2.
XX
KW   Centromere; michrosome; vector; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200055325-A2.
XX
PD   21-SEP-2000.
XX
PF   17-MAR-2000; 2000WO-US007392.
XX
PR   18-MAR-1999;   99US-0125219P.
PR   01-APR-1999;   99US-0127409P.
PR   18-MAY-1999;   99US-0134770P.
PR   13-SEP-1999;   99US-0153584P.
PR   17-SEP-1999;   99US-0154603P.
PR   16-DEC-1999;   99US-0172493P.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Preuss D,  Copenhaver G,  Keith K;
XX
DR   WPI; 2000-587529/55.
XX
PT   Recombinant DNA construct comprising a plant centromere, useful for
PT   producing stably inherited michrosomes which can serve as vectors for the
PT   construction of transgenic plant and animal cells.
XX
PS   Claim 102; Page 321-335; 1449pp; English.
XX
CC   The present invention relates to a recombinant DNA construct of a plant
CC   (Arabidopsis thaliana) centromere. The constructs are useful for
CC   producing stably inherited michrosomes which can serve as vectors for the
CC   construction of transgenic plant and animal cells expressing selected
CC   proteins such as hormones, enzymes, interleukins, clotting factors,
CC   cytokines, antibodies, and growth factors
XX
SQ   Sequence 64415 BP; 18698 A; 13554 C; 13083 G; 18980 T; 0 U; 100 Other;

  Query Match            19.5%;  Score 30.6;  DB 3;  Length 64415;
  Best Local Similarity  58.1%;  Pred. No. 26;
  Matches   54;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy         65 CTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||| | ||        | || |||  | |||| ||||| |  | || |  | | || |
Db       7117 CTTGCTTCTTTAAGTATTATAGTATATTTGCTCCTAAACACTAAACCTAAACCCTACACC 7176

Qy        125 TTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              ||  || | ||  |  | ||| |  ||||| |||
Db       7177 TTAAATCCCAAACCCTAAATCTAATTCCTTAAC 7209


RESULT 14
ADA02717
ID   ADA02717 standard; DNA; 79467 BP.
XX
AC   ADA02717;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Mouse Nfatc1 carcinoma associated gene, SEQ ID NO:1235.
XX
KW   Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW   prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW   gene; ds.
XX
OS   Mus sp.
XX
PN   WO2003057146-A2.
XX
PD   17-JUL-2003.
XX
PF   26-DEC-2002; 2002WO-US041414.
XX
PR   26-DEC-2001; 2001US-00035832.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-587068/55.
XX
PT   New recombinant nucleic acid encoding carcinoma associated protein,
PT   useful for preparing compositions for treating carcinomas.
XX
PS   Claim 1; SEQ ID NO 1235; 245pp; English.
XX
CC   The invention relates to recombinant carcinoma associated (CA) nucleic
CC   acid sequences from mouse and human (ADA01482-ADA03094), and to
CC   recombinant carcinoma associated proteins (CAP) encoded by them. The
CC   invention also encompasses expression vectors and host cells comprising a
CC   CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC   binds to the protein, and a biochip comprising CA nucleic acid or
CC   fragments thereof. The sequences of the invention were identified using
CC   oncogenic retroviruses, which insert into the genome of the host organism
CC   at random. Many of these do not carry transduced host oncogenes or
CC   pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC   direct consequence of the effects of proviral integration into host
CC   protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC   carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC   leukaemia) or a propensity to carcinoma by determination of the sequence
CC   of a CA gene, or by determination of CA gene expression in particular
CC   tissues. CA nucleic acids, proteins and antibodies are also useful as
CC   therapeutic agents and in screening and evaluating drug candidates. The
CC   present sequence represents a specifically claimed murine CA nucleic acid
CC   sequence of the invention. Note: The complete sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;
```

```
  Query Match             19.5%;  Score 30.6;  DB 9;  Length 79467;
  Best Local Similarity   56.4%;  Pred. No. 28;
  Matches   57;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy         51 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 110
              ||||     |  | |||||| |  |||| | ||||  ||  | || || || |||  |
Db      53779 TCCACTAACTGATCCTTGCTTCCTCTGGAACAAGCTGAGTCTTAGTCACACACACACACA 53838

Qy        111 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTC 151
              |  || ||  ||    |||  || ||||   | | ||  ||
Db      53839 CACACACACACACACACCACACACACACAAGCTTGTGGCTC 53879


RESULT 15
ADB72455
ID   ADB72455 standard; DNA; 79467 BP.
XX
AC   ADB72455;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Mouse Nfatc1 gene.
XX
KW   mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW   cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS   Mus sp.
XX
PN   WO2003008583-A2.
XX
PD   30-JAN-2003.
XX
PF   26-DEC-2001; 2001WO-US051291.
XX
PR   02-MAR-2001; 2001US-00798586.
PR   23-OCT-2001; 2001US-00004113.
PR   08-NOV-2001; 2001US-00052482.
PR   30-NOV-2001; 2001US-00997722.
PR   20-DEC-2001; 2001US-00034650.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW,  Engelhard EK;
XX
DR   WPI; 2003-239337/23.
XX
PT   New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT   cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS   Claim 1; SEQ ID NO 283; 2304pp; English.
XX
CC   The invention relates to a novel recombinant nucleic acid comprising a
CC   nucleotide sequence selected from any of the 660 sequences fully defined
CC   in the specification. A polynucleotide of the invention has cytostatic
CC   activity, and may have a use in gene therapy, or in a vaccine. The
CC   recombinant nucleic acids and polypeptides are useful for treating
CC   carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC   sarcomas. The present sequence represents a mouse gene of the invention.
XX
SQ   Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;

  Query Match             19.5%;  Score 30.6;  DB 10;  Length 79467;
  Best Local Similarity   56.4%;  Pred. No. 28;
  Matches   57;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy         51 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 110
              ||||     |  | |||||| |  |||| | ||||  ||  | || || || |||  |
Db      53779 TCCACTAACTGATCCTTGCTTCCTCTGGAACAAGCTGAGTCTTAGTCACACACACACACA 53838

Qy        111 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTC 151
              |  || ||  ||    |||  || ||||   | | ||  ||
Db      53839 CACACACACACACACACCACACACACACAAGCTTGTGGCTC 53879
```

Search completed: March 5, 2006, 18:11:35
Job time : 107.48 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 18:11:59 ; Search time 741.207 Seconds
(without alignments)
9910.279 Million cell updates/sec

Title: US-10-713-381-1_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	157	100.0	687	9	CC656939	CC656939 OGWDQ20TV
	2	157	100.0	915	10	CG224225	CG224225 OG1AG08TV
	3	157	100.0	963	9	CC656933	CC656933 OGWDQ20TM
	4	63	40.1	702	10	CW324514	CW324514 104_819_1
	5	47	29.9	296	10	CW445575	CW445575 fsbb001f1
c	6	37.8	24.1	636	10	CE588104	CE588104 tigr-gss-
c	7	37.8	24.1	643	7	CO689495	CO689495 DG11-25n6
c	8	37.8	24.1	1085	10	AG366773	AG366773 Mus muscu
	9	35.6	22.7	874	10	CZ727957	CZ727957 OC__Ba005
	10	34	21.7	700	2	BB505306	BB505306 BB505306
c	11	33.8	21.5	734	10	CW871670	CW871670 she2h63-4
	12	33.6	21.4	1677	10	CL026582	CL026582 CH216-23G
	13	32.8	20.9	440	1	AI744861	AI744861 tr16b06.x
c	14	32.6	20.8	648	8	CX446438	CX446438 JGI_XZG22
c	15	32.6	20.8	740	10	CW558661	CW558661 OA_ABa008
c	16	32.6	20.8	758	10	CW597425	CW597425 OA_ABa013
	17	32.6	20.8	775	8	CX424281	CX424281 JGI_XZG64
	18	32.6	20.8	804	7	CR575123	CR575123 CR575123
c	19	32.6	20.8	843	10	CW599444	CW599444 OA_ABa013
	20	32.6	20.8	864	8	CX404228	CX404228 JGI_XZT35
	21	32.6	20.8	868	8	DN096073	DN096073 JGI_CABE6
c	22	32.6	20.8	1101	10	CNS003FV	AL064657 Drosophil
	23	32.4	20.6	347	2	BE690384	BE690384 uw66f12.y
c	24	32.4	20.6	483	1	AJ746802	AJ746802 AJ746802
c	25	32.4	20.6	497	1	AW990724	AW990724 uf10b11.x
c	26	32.4	20.6	498	2	BE632212	BE632212 uu11f01.x
c	27	32.4	20.6	500	2	BF225639	BF225639 uy44e02.x
	28	32.4	20.6	527	2	BE448824	BE448824 ut88c10.y
	29	32.4	20.6	528	7	CN243061	CN243061 EST008934
	30	32.4	20.6	539	5	BX528623	BX528623 BX528623
	31	32.4	20.6	542	1	AA673664	AA673664 vo57h06.r
c	32	32.4	20.6	543	2	BE133867	BE133867 ug13b03.x
c	33	32.4	20.6	545	2	BF020472	BF020472 uw66f12.x
	34	32.4	20.6	547	5	BX517541	BX517541 BX517541
c	35	32.4	20.6	570	2	BE134144	BE134144 ug16f12.x
	36	32.4	20.6	621	1	BB004373	BB004373 BB004373
	37	32.4	20.6	742	3	BI453933	BI453933 603174018
	38	32.4	20.6	859	10	CNS004YY	AL055406 Drosophil
c	39	32.4	20.6	914	2	BF682258	BF682258 602117694
c	40	32.4	20.6	1039	11	CNS055LA	AL322183 Tetraodon
	41	32.4	20.6	2194	4	BC048853	BC048853 Mus muscu
	42	32.4	20.6	2454	4	AK028832	AK028832 Mus muscu
	43	32.2	20.5	432	9	BH752563	BH752563 SALK_0188
	44	32.2	20.5	960	10	DU054394	DU054394 91884 Tom
c	45	32.2	20.5	979	10	CZ954687	CZ954687 268372 To

ALIGNMENTS

RESULT 1
CC656939/c
LOCUS CC656939 687 bp DNA linear GSS 19-JUN-2003
DEFINITION OGWDQ20TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15, genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 687)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGWDQ20TM
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: methylation filtered.
FEATURES Location/Qualifiers
source 1. .687
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0554D15"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 100.0%; Score 157; DB 9; Length 687;
Best Local Similarity 100.0%; Pred. No. 2.2e-39;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 362

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 302

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
                |||||||||||||||||||||||||||||||||||
Db        301 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 265


RESULT 2
CG224225
LOCUS       CG224225                915 bp    DNA     linear   GSS 22-AUG-2003
DEFINITION  OG1AG08TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0716B15,
            genomic survey sequence.
ACCESSION   CG224225
VERSION     CG224225.1  GI:34124113
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1AG08TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0716B15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             100.0%;  Score 157;  DB 10;  Length 915;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-39;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 593

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 653

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
                |||||||||||||||||||||||||||||||||||
Db        654 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 690


RESULT 3
CC656933
LOCUS       CC656933                963 bp    DNA     linear   GSS 19-JUN-2003
DEFINITION  OGWDQ20TM ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15,
            genomic survey sequence.
ACCESSION   CC656933
VERSION     CC656933.1  GI:32060225
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 963)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OGWDQ20TV
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .963
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0554D15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             100.0%;  Score 157;  DB 9;  Length 963;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-39;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        752 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 811

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        812 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 871

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
                |||||||||||||||||||||||||||||||||||
Db        872 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 908


RESULT 4
CW324514
LOCUS       CW324514                702 bp    DNA     linear   GSS 31-OCT-2004
DEFINITION  104_819_11477203_148_35910_078 Sorghum methylation filtered library
            (LibID: 104) Sorghum bicolor genomic clone 11477203, genomic survey
            sequence.
ACCESSION   CW324514
VERSION     CW324514.1  GI:55040702
KEYWORDS    GSS.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
            Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
            McMenamy,J., Smith,M., Holeman,H., Roe,B.A, Wiley,G., Korf,I.F.,
            Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
            Martienssen,R.A.
```

```
  TITLE     Sorghum genome sequencing by methylation filtration
  JOURNAL   PLoS Biol. 3 (1), e13 (2005)
   PUBMED   15660154
COMMENT     Contact: Bedell JA
            Orion Genomics, LLC
            4041 Forest Park Ave, St. Louis, MO 63108, USA
            Tel: 314 615 6979
            Fax: 314 615 5975
            Email: jbedell@oriongenomics.com
            Plate: 819  row: c  column: 19
            Seq primer: SWfor Forward
            Class: methylation filtered
            High quality sequence stop: 702.
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="Sorghum bicolor"
                     /mol_type="genomic DNA"
                     /cultivar="ATx623"
                     /db_xref="taxon:4558"
                     /clone="11477203"
                     /clone_lib="Sorghum methylation filtered library (LibID:
                     104)"
                     /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
                     prepared from purified nuclei was randomly sheared,
                     end-repaired, size fractionated to enrich for the 0.5 to 5
                     kb fraction, ligated into HincII-digested pBCSK(-) vector
                     and electroporated into E. coli cells. This is a
                     methylation filtered library."
ORIGIN

  Query Match          40.1%;  Score 63;  DB 10;  Length 702;
  Best Local Similarity   93.0%;  Pred. No. 3.9e-09;
  Matches   66;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 123
              || || |||||||||||||||||| |||||||||||||||||||| |||| |||||||
Db         53 CCCTGGCTATTCTGAACCAAGAGGACACCTACTCCCAAACAATCCATGTTACCCATGCAA 112

Qy        124 CTTCCATGCAA 134
              ||||||||||
Db        113 CTTCCATGCAA 123


RESULT 5
CW445575
LOCUS       CW445575                296 bp    DNA     linear   GSS 02-NOV-2004
DEFINITION  fsbb001f170m16k0 Sorghum methylation filtered library (LibID: 104)
            Sorghum bicolor genomic clone fsbb001f170m16, genomic survey
            sequence.
ACCESSION   CW445575
VERSION     CW445575.1  GI:55193536
KEYWORDS    GSS.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 296)
  AUTHORS   Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D.,
            Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K.,
            McMenamy,J., Smith,M., Holeman,H., Roe,B.A, Wiley,G., Korf,I.F.,
            Rabinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,J.A. and
            Martienssen,R.A.
  TITLE     Sorghum genome sequencing by methylation filtration
  JOURNAL   PLoS Biol. 3 (1), e13 (2005)
   PUBMED   15660154
COMMENT     Contact: Bedell JA
            Orion Genomics, LLC
            4041 Forest Park Ave, St. Louis, MO 63108, USA
            Tel: 314 615 6979
            Fax: 314 615 5975
            Email: jbedell@oriongenomics.com
            Plate: fsbb001f170  row: m  column: 16
            Seq primer: k Reverse
            Class: methylation filtered
            High quality sequence stop: 296.
FEATURES             Location/Qualifiers
     source          1. .296
                     /organism="Sorghum bicolor"
                     /mol_type="genomic DNA"
                     /cultivar="ATx623"
                     /db_xref="taxon:4558"
                     /clone="fsbb001f170m16"
                     /clone_lib="Sorghum methylation filtered library (LibID:
                     104)"
                     /note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
                     prepared from purified nuclei was randomly sheared,
                     end-repaired, size fractionated to enrich for the 0.5 to 5
                     kb fraction, ligated into HincII-digested pBCSK(-) vector
                     and electroporated into E. coli cells. This is a
                     methylation filtered library."
ORIGIN

  Query Match          29.9%;  Score 47;  DB 10;  Length 296;
  Best Local Similarity   85.3%;  Pred. No. 0.00042;
  Matches   64;  Conservative    0;  Mismatches   10;  Indels    1;  Gaps    1;

Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT-CCATCTTACTCATGCA 122
              || || |||||||||||||||||| |||||||||||||||||| |||| |||| ||||||
Db        221 CCCTGGCTATTCTGAACCAAGAGGACACCTACTCCCAAACAATCCCATGTTACCCATGCA 280

Qy        123 ACTTCCATGCAAACA 137
              ||||||    ||||
Db        281 ACTTCCCATGCAACA 295


RESULT 6
CE588104/c
LOCUS       CE588104                636 bp    DNA     linear   GSS 28-SEP-2003
DEFINITION  tigr-gss-dog-17000366359543 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE588104
VERSION     CE588104.1  GI:36904885
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE     The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
   PUBMED   14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site_1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

  Query Match          24.1%;  Score 37.8;  DB 10;  Length 636;
```

```
  Best Local Similarity   58.4%;  Pred. No. 0.48;
  Matches   66;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          45 TCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACA 104
               | || || |   || |||| | | | ||  || ||||| | ||||||  |    || |
Db         292 TATTTTTTCCCAGTTCAAGGCATCCTTGTTTGGACCCAAGGGTATACCTCATTTTTAAAA 233

Qy         105 ATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
               || ||| ||| |   | || || ||  || | |   | |   || |  |||||
Db         232 ATACATTTTATTATAGAAATTTTCAAACACAAAAATAAAATGGTGTAGTGAAC 180


RESULT 7
CO689495/c
LOCUS       CO689495                643 bp    mRNA    linear   EST 26-JUL-2004
DEFINITION  DG11-25n6 DG11-kidney Canis familiaris cDNA 3', mRNA sequence.
ACCESSION   CO689495
VERSION     CO689495.1  GI:50638161
KEYWORDS    EST.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 643)
  AUTHORS   Schlueter,T., Hermanns,J., Weindel,M., Schuette,D., Kranz,H.,
            Henrich,J. and Loebbert,R.
  TITLE     Dog arrayTAG cDNA clone collection
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Thomas Schlueter
            LION bioscience AG
            Waldhoferstrasse 98, D-69123 Heidelberg, Germany
            Tel: +49 6221 4038 150
            Fax: +49 6221 4038 290
            Email: Thomas.Schlueter@lionbioscience.com.
FEATURES             Location/Qualifiers
     source          1. .643
                     /organism="Canis familiaris"
                     /mol_type="mRNA"
                     /strain="Beagle"
                     /db_xref="taxon:9615"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="DG11-kidney"
                     /note="Organ: kidney; Vector: Dog pBluescript LION"
ORIGIN

  Query Match             24.1%;  Score 37.8;  DB 7;  Length 643;
  Best Local Similarity   58.4%;  Pred. No. 0.49;
  Matches   66;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          45 TCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACA 104
               | || || |   || |||| | | | ||  || ||||| | ||||||  |    || |
Db         333 TATTTTTTCCCAGTTCAAGGCATCCTTGTTTGGACCCAAGGGTATACCTCATTTTTAAAA 274

Qy         105 ATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
               || ||| ||| |   | || || ||  || | |   | |   || |  |||||
Db         273 ATACATTTTATTATAGAAATTTTCAAACACAAAAATAAAATGGTGTAGTGAAC 221


RESULT 8
AG366773/c
LOCUS       AG366773               1085 bp    DNA     linear   GSS 21-DEC-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-170B12.TJ, genomic survey
            sequence.
ACCESSION   AG366773
VERSION     AG366773.1  GI:47977978
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus (Japanese wild mouse)
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
            Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and
            Shiroishi,T.
  TITLE     Contribution of Asian mouse subspecies Mus musculus molossinus to
            genomic constitution of strain C57BL/6J, as defined by BAC-end
            sequence-SNP analysis
  JOURNAL   Genome Res. 14 (12), 2439-2447 (2004)
   PUBMED   15574823
REFERENCE   2  (bases 1 to 1085)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .1085
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-170B12.TJ"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             24.1%;  Score 37.8;  DB 10;  Length 1085;
  Best Local Similarity   51.9%;  Pred. No. 0.57;
  Matches   69;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy          22 CTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACC 81
               || |   | |  ||  ||   |  ||| | || || ||| | |||  ||| ||||  ||
Db         719 CTCCCCTCCTCCTCCCCCTCCCTCCTTCTCCCCTCCTCCTACCCTCTCCTTTTCTATACT 660

Qy          82 AAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCA 141
               |     ||| |  | ||| |  || |||    |  | ||||  ||     ||  |   |
Db         659 ACATNAATAACNCCCCCCNANNAAACCANNANAACCCCGCAATTTTTTNANAANNATAAA 600

Qy         142 CATATGTTTCCTG 154
               || |  |||  ||
Db         599 CAAACTTTTTTG 587


RESULT 9
CZ727957
LOCUS       CZ727957                874 bp    DNA     linear   GSS 25-JUL-2005
DEFINITION  OC__Ba0055L24.f OC__Ba Oryza coarctata genomic clone OC__Ba0055L24
            5', genomic survey sequence.
ACCESSION   CZ727957
VERSION     CZ727957.1  GI:71136431
KEYWORDS    GSS.
SOURCE      Oryza coarctata (Porteresia coarctata)
  ORGANISM  Oryza coarctata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
```

```
REFERENCE   1  (bases 1 to 874)
  AUTHORS   Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
  TITLE     OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0055  row: L  column: 24
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .874
                     /organism="Oryza coarctata"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:77588"
                     /clone="OC__Ba0055L24"
                     /tissue_type="leaves"
                     /dev_stage="mature"
                     /lab_host="DH10B"
                     /clone_lib="OC__Ba"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match           22.7%;  Score 35.6;  DB 10;  Length 874;
  Best Local Similarity  78.6%;  Pred. No. 2.7;
  Matches   55;  Conservative    0;  Mismatches   14;  Indels    1;  Gaps    1;

Qy         64 CCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAA 123
              ||||| |||||| ||||||| ||| |||||||| |||| ||| |  ||||||||||
Db        778 CCTTGGCTATTCCGAACCAACAGG-CACCTACTCTCAAAACATCACGCCGACTCATGCAA 836

Qy        124 CTTCCATGCA 133
              | ||||||
Db        837 CCTCCATGTA 846


RESULT 10
BB505306
LOCUS       BB505306                700 bp    mRNA    linear   EST 25-OCT-2001
DEFINITION  BB505306 RIKEN full-length enriched, 10 days lactation, adult
            female mammary gland Mus musculus cDNA clone D730002G06, mRNA
            sequence.
ACCESSION   BB505306
VERSION     BB505306.2  GI:16442791
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     On Jul 27, 2000 this sequence version replaced gi:9514268.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute.
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
            Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to.
            prepare mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="D730002G06"
                     /sex="female"
                     /tissue_type="mammary gland"
                     /dev_stage="10 days lactation, adult"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 10 days lactation,
                     adult female mammary gland"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                     was cleaved with BamHI and XhoI. Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC I."
ORIGIN

  Query Match           21.7%;  Score 34;  DB 2;  Length 700;
  Best Local Similarity  54.9%;  Pred. No. 8.3;
  Matches   67;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;

Qy         27 TGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAG 86
              | ||| |||  || | | |||  || | | || |     ||| | | ||       |||
Db        195 TTCTTTTTCTCCCCTCCCTCTGATTTCTTTGTACCTCACTTTCATTTTGGCCTTTAAGTT 254

Qy         87 GATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATAT 146
              | |    | ||||||| || |||| || | |       |   ||| | ||||| |  | ||
```

```
Db        255 GTTCTAGATTCCCAAAGAAGCCATTTTGCACTACAGTCAAACATACTAACACCCTGAAAT 314

Qy        147 GT 148
              ||
Db        315 GT 316


RESULT 11
CW871670/c
LOCUS       CW871670               734 bp    DNA     linear   GSS 12-FEB-2005
DEFINITION  she2h63-44.g_013.ab1 Whole-genome shotgun library of the elephant
            shark (aka elephant fish) Callorhinchus milii genomic, genomic
            survey sequence.
ACCESSION   CW871670
VERSION     CW871670.1  GI:59698305
KEYWORDS    GSS.
SOURCE      Callorhinchus milii (elephantfish)
  ORGANISM  Callorhinchus milii
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
            Holocephali; Chimaeriformes; Callorhinchidae; Callorhinchus.
REFERENCE   1  (bases 1 to 734)
  AUTHORS   Venkatesh,B., Tay,A., Dandona,N., Patil,J.G. and Brenner,S.
  TITLE     A compact cartilaginous fish model genome
  JOURNAL   Curr. Biol. 15 (3), R82-R83 (2005)
   PUBMED   15694293
COMMENT     Contact: Venkatesh B
            Molecular Genetics Lab
            Institute of Molecular and Cell Biology
            61 Biopolis Drive, Singapore 138673
            Tel: 65 6586 9571
            Fax: 65 6779 1117
            Email: mcbbv@imcb.a-star.edu.sg
            Whole-genome shotgun sequences of the elephant shark (aka elephant
            fish)
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .734
                     /organism="Callorhinchus milii"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7868"
                     /sex="Male"
                     /tissue_type="Testis"
                     /clone_lib="Whole-genome shotgun library of the elephant
                     shark (aka elephant fish)"
ORIGIN

  Query Match             21.5%;  Score 33.8;  DB 10;  Length 734;
  Best Local Similarity   54.4%;  Pred. No. 9.8;
  Matches   68;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              |||| ||| |||  | |||  | || || || |   | ||  | | |||||  ||   ||
Db        238 CGTGGCATTTCAATTTGCACTCCACGTGATTATGGCATCGAACTTGTTGTTAAATGACCC 179

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              | |    ||        ||       || || ||| |||| |||  |||| | || ||
Db        178 GAACTACACCACCCACAGACACCACAGACACATACACCCACACACACCATTTAACACAGA 119

Qy        121 CAACT 125
              | ||
Db        118 AATCT 114


RESULT 12
CL026582
LOCUS       CL026582               1677 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-23G24_Sp6.1 CH216 Xenopus tropicalis genomic clone
            CH216-23G24, genomic survey sequence.
ACCESSION   CL026582
VERSION     CL026582.1  GI:40470443
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1677)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 125
            High quality sequence stop: 217.
FEATURES             Location/Qualifiers
     source          1. .1677
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-23G24"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             21.4%;  Score 33.6;  DB 10;  Length 1677;
  Best Local Similarity   53.9%;  Pred. No. 15;
  Matches   69;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy          3 TGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAA 62
              |||| |||||||||    |    | || || ||   |  |||  | |||  |
Db        769 TGTCGTCTCACATGCACTTACTCCTGATTCTTTTGCTCTTCTACATGTGTCCATCCGGTT 828

Qy         63 GCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCA 122
              | |   ||||||| |     | |  |||| ||| ||     |  || |||  |||  ||
Db        829 TCATCAACTATTCTTACAGTTCATGTCACCTCCTCTCACTTTCTATATATTATTCAGACA 888

Qy        123 ACTTCCAT 130
               |||||||
Db        889 TCTTCCAT 896


RESULT 13
AI744861
LOCUS       AI744861               440 bp    mRNA    linear   EST 21-JUN-1999
DEFINITION  tr16b06.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2218451 3',
            mRNA sequence.
ACCESSION   AI744861
VERSION     AI744861.1  GI:5113149
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
```

```
                     /db_xref="taxon:4532"
                     /clone="OA_ABa0080B07"
                     /tissue_type="young leaves"
                     /lab_host="DH10B T1 phage resistant"
                     /clone_lib="OA_ABa"
                     /note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             20.8%;  Score 32.6;  DB 10;  Length 740;
  Best Local Similarity   63.3%;  Pred. No. 24;
  Matches   50;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy          79 ACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACAC 138
               ||||| |   |  |||||||| ||||||| ||||||  | |  ||| |   | ||||  |
Db         653 ACCAATATACTTACTACTCCCCAACAATGCATCTTCTTAAGCCAAATGGAAGGCAACAAT 594

Qy         139 GCACATATGTTTCCTGAAC 157
               |||  |  ||   |||| |
Db         593 GCAACTGCGTGAGCTGAGC 575


Search completed: March  6, 2006, 01:57:42
Job time : 745.207 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 21:54:12 ; Search time 43.8947 Seconds
(without alignments)
6357.883 Million cell updates/sec

Title: US-10-713-381-1_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	157	100.0	1394	3	US-08-880-499-1	Sequence 1, Appli
	2	157	100.0	1394	3	US-08-880-499-2	Sequence 2, Appli
c	3	30.4	19.4	385136	3	US-09-949-016-16073	Sequence 16073, A
c	4	29.6	18.9	1185	3	US-09-543-681A-1081	Sequence 1081, Ap
c	5	29.2	18.6	1830121	3	US-09-557-884-1	Sequence 1, Appli
c	6	29.2	18.6	1830121	3	US-09-643-990A-1	Sequence 1, Appli
c	7	29.2	18.6	1830121	3	US-10-158-865-1	Sequence 1, Appli
c	8	28.8	18.3	64190	3	US-09-949-016-14712	Sequence 14712, A
c	9	28.8	18.3	64190	3	US-09-949-016-14713	Sequence 14713, A
	10	28	17.8	11022	3	US-09-949-016-15502	Sequence 15502, A
	11	28	17.8	41815	3	US-09-949-016-17447	Sequence 17447, A
c	12	27.8	17.7	601	3	US-09-949-016-191201	Sequence 191201,
c	13	27.8	17.7	11490	3	US-09-949-016-14149	Sequence 14149, A
	14	27.8	17.7	14952	3	US-09-949-016-16740	Sequence 16740, A
	15	27.6	17.6	927	3	US-09-937-862B-37	Sequence 37, Appl
c	16	27.4	17.5	1413	3	US-09-710-279-1813	Sequence 1813, Ap
c	17	27.4	17.5	1428	3	US-09-134-001C-1413	Sequence 1413, Ap
c	18	27.4	17.5	1794	3	US-09-270-767-10489	Sequence 10489, A
c	19	27.4	17.5	2472	2	US-08-425-299A-2	Sequence 2, Appli
	20	27.4	17.5	2926	3	US-09-710-279-3747	Sequence 3747, Ap
c	21	27.4	17.5	3014	3	US-09-710-279-3680	Sequence 3680, Ap
	22	27.4	17.5	3599	3	US-09-710-279-4003	Sequence 4003, Ap
c	23	27.2	17.3	3445	3	US-09-949-016-2556	Sequence 2556, Ap
c	24	27.2	17.3	3588	3	US-09-566-921-23	Sequence 23, Appl
c	25	27.2	17.3	6751	2	US-07-882-925A-5	Sequence 5, Appli
c	26	27.2	17.3	6751	2	US-08-184-012C-5	Sequence 5, Appli
c	27	27.2	17.3	8939	3	US-09-949-016-12703	Sequence 12703, A
c	28	27.2	17.3	9076	3	US-09-949-016-17368	Sequence 17368, A
	29	27.2	17.3	48994	3	US-09-949-016-14091	Sequence 14091, A
	30	27.2	17.3	60593	3	US-09-949-016-13779	Sequence 13779, A
c	31	27.2	17.3	62776	3	US-09-949-016-17576	Sequence 17576, A
c	32	27.2	17.3	100848	3	US-09-596-002-39	Sequence 39, Appl
c	33	27.2	17.3	158735	3	US-09-949-016-11989	Sequence 11989, A
c	34	27.2	17.3	158735	3	US-09-949-016-17130	Sequence 17130, A
c	35	27	17.2	879	3	US-09-134-001C-678	Sequence 678, App
	36	27	17.2	44180	3	US-09-949-016-12088	Sequence 12088, A
	37	27	17.2	47030	3	US-09-949-016-13037	Sequence 13037, A
	38	27	17.2	47030	3	US-09-949-016-13038	Sequence 13038, A
	39	27	17.2	47030	3	US-09-949-016-15039	Sequence 15039, A
	40	27	17.2	47030	3	US-09-949-016-15040	Sequence 15040, A
c	41	27	17.2	160552	3	US-09-593-828-11	Sequence 11, Appl
	42	27	17.2	194537	3	US-09-949-016-12928	Sequence 12928, A
	43	27	17.2	201529	3	US-09-949-016-12740	Sequence 12740, A
	44	26.8	17.1	265	3	US-09-513-999C-8379	Sequence 8379, Ap
c	45	26.8	17.1	271	3	US-08-946-026-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-880-499-1
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-880-499-1

Query Match 100.0%; Score 157; DB 3; Length 1394;

```
  Best Local Similarity   100.0%;  Pred. No. 9.4e-44;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 2
US-08-880-499-2
; Sequence 2, Application US/08880499
; Patent No. 6037523
;  GENERAL INFORMATION:
;    APPLICANT:  Albertson, Marc C.
;    APPLICANT:  Fox, Tim W.
;    APPLICANT:  Carl, Garnaat W.
;    APPLICANT:  Huffman, Gary A.
;    APPLICANT:  Kendall, Timmy L.
;    TITLE OF INVENTION:  MALE TISSUE-PREFERRED REGULATORY REGION
;    TITLE OF INVENTION:  AND METHOD OF USING SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PIONEER HI-BRED INTERNATIONAL, INC.
;      STREET:  Darwin Building, 7100 N.W. 62nd Ave., P.O.
;      STREET:  Box 1000
;      CITY:  Johnston
;      STATE:  Iowa
;      COUNTRY:  USA
;      ZIP:  50131
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/880,499
;      FILING DATE:  CONCURRENTLY HEREWITH
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sweeney, Patricia A.
;      REGISTRATION NUMBER:  32,733
;      REFERENCE/DOCKET NUMBER:  0578
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (515) 248-4800
;      TELEFAX:  (515) 248-4844
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1394 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-880-499-2

  Query Match             100.0%;  Score 157;  DB 3;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 9.4e-44;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 3
US-09-949-016-16073/c
; Sequence 16073, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16073
;   LENGTH: 385136
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(385136)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16073

  Query Match             19.4%;  Score 30.4;  DB 3;  Length 385136;
  Best Local Similarity   57.3%;  Pred. No. 14;
  Matches   55;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy         55 TCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTA 114
              || ||||   |  |||||| |||| |     |  | ||| | ||||||  |     ||
Db     253908 TCCTCCAGATCGCTCCTATTTTGAAATATATAAACGCTTACACACAAACACACTCATATA 253849

Qy        115 CTCATGCAACTTCCATGCAAACACGCACATATGTTT 150
              |||   ||| |||    | ||| ||||| | |||
Db     253848 TCCATTTTACTCCCACCACACCACACACATTTTTTT 253813


RESULT 4
US-09-543-681A-1081/c
; Sequence 1081, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1081
;   LENGTH: 1185
;   TYPE: DNA
;   ORGANISM: Proteus mirabilis
US-09-543-681A-1081

  Query Match             18.9%;  Score 29.6;  DB 3;  Length 1185;
  Best Local Similarity   59.5%;  Pred. No. 3.4;
  Matches   50;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy         35 CAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTA 94
```

```
Db        940 CAACAATTTGTTTTTGTCCACCAGAAAGACCTAACCCATTTTCTCCCAAGGACATATCTA 881

Qy         95 CTCCCAAACAATCCATCTTACTCA 118
              |||| | || | | ||| | |
Db        880 ATCCCCTAGGATGCTTTTTAATAA 857


RESULT 5
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
;   GENERAL INFORMATION:
;        APPLICANT: Fleischmann et al.
;        TITLE OF INVENTION: The Nucleotide sequence of
;                            the Haemophilus influenzae Rd Genome, Fragments
;                            Thereof, and Uses Thereof
;        NUMBER OF SEQUENCES: 1
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3 1/2 inch diskette
;             COMPUTER: Dell Pentium
;             OPERATING SYSTEM: MS DOS v6.22
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/557,884
;             FILING DATE: 25-Apr-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/476,102
;             FILING DATE: JUN-5-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Michelle S. Marks
;             REGISTRATION NUMBER: 41,971
;             REFERENCE/DOCKET NUMBER: PB186P3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-309-8504
;             TELEFAX: 301-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1830121 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

  Query Match           18.6%;  Score 29.2;  DB 3;  Length 1830121;
  Best Local Similarity  51.5%;  Pred. No. 63;
  Matches   67;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy         12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
              || |   || |  | ||||||  |    | | ||          | |||| | ||||
Db     677275 ACTTTAAATTCAGCCTGCTTGCTTGCTCTATTGCATTTGCATTAAATGCAAGCATAGCCT 677216

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              || |||  | |  |    ||| ||  | ||| | ||| |  ||  |  | |  |||
Db     677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156

Qy        132 CAAACACGCA 141
              ||| ||  ||
Db     677155 CAACCAACCA 677146


RESULT 6
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
;   GENERAL INFORMATION:
;        APPLICANT: Robert D. Fleischmann
;                   Mark D. Adams
;                   Owen White
;                   Hamilton O. Smith
;                   J. Craig Venter
;        TITLE OF INVENTION: The Nucleotide sequence of
;                            the Haemophilus influenzae Rd Genome, Fragments
;                            Thereof, and Uses Thereof
;        NUMBER OF SEQUENCES: 1
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville,
;             STATE: MD
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3 1/2 inch diskette
;             COMPUTER: Dell Pentium
;             OPERATING SYSTEM: MS DOS v6.22
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/643,990A
;             FILING DATE: 23-Aug-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/487,429
;             FILING DATE: 1995-06-07
;             APPLICATION NUMBER: 08/426,787
;             FILING DATE: 1995-04-21
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB186P1C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 301-610-5790
;             TELEFAX: 310-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1830121 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

  Query Match           18.6%;  Score 29.2;  DB 3;  Length 1830121;
  Best Local Similarity  51.5%;  Pred. No. 63;
  Matches   67;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy         12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
              || |   || |  | ||||||  |    | | ||          | |||| | ||||
Db     677275 ACTTTAAATTCAGCCTGCTTGCTTGCTCTATTGCATTTGCATTAAATGCAAGCATAGCCT 677216

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              || |||  | |  |    ||| ||  | ||| | ||| |  ||  |  | |  |||
Db     677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156

Qy        132 CAAACACGCA 141
              ||| ||  ||
Db     677155 CAACCAACCA 677146


RESULT 7
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fragn
; Patent No. 6846651
```

; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature

```
;   LOCATION: (122336)..(122336)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131340)..(131340)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (131360)..(131360)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (139910)..(139910)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (140398)..(140398)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (142750)..(142750)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145058)..(145058)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145171)..(145171)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145942)..(145942)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (147197)..(147197)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (150841)..(150841)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (152500)..(152500)
;   OTHER INFORMATION: n equals a,t,c, or g
;   FEATURE:
;   NAME/KEY: misc_feature

  Query Match           18.6%;  Score 29.2;  DB 3;  Length 1830121;
  Best Local Similarity   51.5%;  Pred. No. 63;
  Matches   67;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy         12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
              ||  |   ||  |  | |||||| |     |  | |  ||          | ||||| | ||||
Db     677275 ACTTTAAATTCAGCCTGCTTGCTTGCTCTATTGCATTTGCATTAAATGCAAGCATAGCCT 677216

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              || |||  | |  |   ||| ||   | ||| | ||| |  ||  |   | |  |||
Db     677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACAAC 677156

Qy        132 CAAACACGCA 141
              ||| ||  ||
Db     677155 CAACCAACCA 677146


RESULT 8
US-09-949-016-14712/c
; Sequence 14712, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14712
;   LENGTH: 64190
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14712

  Query Match           18.3%;  Score 28.8;  DB 3;  Length 64190;
  Best Local Similarity   62.5%;  Pred. No. 27;
  Matches   45;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy         55 TCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTA 114
              || |||||   || ||  | || | ||| || |  ||  | | | || ||| |||   |
Db        355 TCATCCAATAGTTTCCACTACTAATCCATGATGGAACTGAGTACTAATCAAGCCAATATT 296

Qy        115 CTCATGCAACTT 126
              |||||| ||| |
Db        295 CTCATGTAACAT 284


RESULT 9
US-09-949-016-14713/c
; Sequence 14713, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14713
;   LENGTH: 64190
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14713

  Query Match           18.3%;  Score 28.8;  DB 3;  Length 64190;
  Best Local Similarity   62.5%;  Pred. No. 27;
  Matches   45;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy         55 TCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTA 114
              || |||||   || ||  | || | ||| || |  ||  | | | || ||| |||   |
Db        355 TCATCCAATAGTTTCCACTACTAATCCATGATGGAACTGAGTACTAATCAAGCCAATATT 296

Qy        115 CTCATGCAACTT 126
              |||||| ||| |
Db        295 CTCATGTAACAT 284


RESULT 10
US-09-949-016-15502
; Sequence 15502, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15502
;   LENGTH: 11022
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-15502

  Query Match             17.8%;  Score 28;  DB 3;  Length 11022;
  Best Local Similarity   63.2%;  Pred. No. 27;
  Matches   43;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         15 TGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATT 74
              |||| || | |||| |  ||  |  ||| | |||| ||||     ||||||| | |
Db       4432 TGGCTTATTTCATGGTGTTTTTAGGGTTAATGTTGTAGCATCCATTTAGCCTTGTCAAGG 4491

Qy         75 CTGAACCA 82
              ||||| ||
Db       4492 CTGAATCA 4499


RESULT 11
US-09-949-016-17447
; Sequence 17447, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17447
;   LENGTH: 41815
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17447

  Query Match             17.8%;  Score 28;  DB 3;  Length 41815;
  Best Local Similarity   58.3%;  Pred. No. 43;
  Matches   49;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy         62 AGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGC 121
              || | |||  |||| |  ||  | ||| || |||| |  |  ||  | |||||| |
Db      33058 AGACATGCAGATTCACACACATCCGCATAGCTGCTCCTACGGATCCCTGGTCACTCATTC 33117

Qy        122 AACTTCCATGCAAACACGCACATA 145
              |  |  |||  | ||| |||  ||
Db      33118 ATGTCACATAGAGACATGCATGTA 33141


RESULT 12
US-09-949-016-191201/c
; Sequence 191201, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 191201
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-191201

  Query Match             17.7%;  Score 27.8;  DB 3;  Length 601;
  Best Local Similarity   58.0%;  Pred. No. 11;
  Matches   47;  Conservative    1;  Mismatches   33;  Indels    0;  Gaps    0;

Qy         57 GTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACT 116
              ||||   | |||| |   |  :  || | | | |   || | || ||| ||  ||| | |
Db        322 GTCCTGACGTTGCATGCACACRTGCACGCGCACAAACACACACACACACTCTCTCTCATT 263

Qy        117 CATGCAACTTCCATGCAAACA 137
              ||||||| |   ||| ||| |
Db        262 CATGCAAATAGGATGGAAATA 242


RESULT 13
US-09-949-016-14149/c
; Sequence 14149, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14149
;   LENGTH: 11490
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-14149

  Query Match             17.7%;  Score 27.8;  DB 3;  Length 11490;
  Best Local Similarity   52.1%;  Pred. No. 32;
  Matches   62;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy         38 CCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTC 97
              | |  | |||  |||| |   ||   || ||||     |   |||| |||||  || || |
Db       5642 CAGGGCCTCTGCTTCCCTGCCCCTCCCCCTGCCCCCACCTCACCAGGAGGAAGCCCACGC 5583

Qy         98 CCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAA 156
              ||   ||  ||   |  | ||| |   |      |||  |  |   | |||||  ||  ||
Db       5582 CCTTGCACACCTGATGCCACATCCGCTTGAGTGGCAGGCGAGTGAAGATGTTGCCCAAA 5524
```

```
RESULT 14
US-09-949-016-16740
; Sequence 16740, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16740
;   LENGTH: 14952
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16740

  Query Match             17.7%;  Score 27.8;  DB 3;  Length 14952;
  Best Local Similarity   59.5%;  Pred. No. 35;
  Matches   47;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          67 TGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTT 126
               || | |   ||| |||| || |||  || | |  |  ||| | |||  | || | ||||
Db          70 TGTCAACCGTGATCCAACAGTATAATTAGTTCTTAGGAATTCCTCTCTCACAGGTAACTA 129

Qy         127 CCATGCAAACACGCACATA 145
                 |  | |||| |||| |
Db         130 GGTTTTACACACACACACA 148


RESULT 15
US-09-937-862B-37
; Sequence 37, Application US/09937862B
; Patent No. 6846621
; GENERAL INFORMATION:
;  APPLICANT: Oberste, M. Steven
;  APPLICANT:  Maher, Kaija
;  APPLICANT:  Kilpatrick, David R.
;  APPLICANT:  Pallansch, Mark A.
;  TITLE OF INVENTION: TYPING OF HUMAN ENTEROVIRUSES
;  FILE REFERENCE: 14114.0353U2
;  CURRENT APPLICATION NUMBER: US/09/937,862B
;  CURRENT FILING DATE:  2001-09-28
;  PRIOR APPLICATION NUMBER: PCT/US00/07828
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/127,464
;  PRIOR FILING DATE: 1999-03-31
;  NUMBER OF SEQ ID NOS: 89
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 37
;   LENGTH: 927
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence; No. 6846621e =
;   OTHER INFORMATION: synthetic construct
US-09-937-862B-37

  Query Match             17.6%;  Score 27.6;  DB 3;  Length 927;
  Best Local Similarity   55.1%;  Pred. No. 15;
  Matches   54;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          53 CATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCT 112
                    |||| || || || |   ||||||  ||| |  | | | |  | |||| |   || |||
Db         551 CATCATCAAACCCCTCAATATTCTACACCTATGGAACAGCACCACCCAGAATTTCGATCC 610

Qy         113 TACTCATGCAACTTCCATGCAAACACGCACATATGTTT 150
                 |  ||    || ||  |    || ||||| | | |
Db         611 CTTTTGTGGGCATTACAAACGCGTACTCACATTTTTAT 648


Search completed: March  5, 2006, 22:36:21
Job time : 48.8947 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 23:29:31 ; Search time 225.751 Seconds
(without alignments)
5750.985 Million cell updates/sec

Title: US-10-713-381-1_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	157	100.0	255	8	US-10-713-381-9	Sequence 9, Appli
	2	157	100.0	1394	8	US-10-713-381-1	Sequence 1, Appli
	3	157	100.0	1394	8	US-10-713-381-2	Sequence 2, Appli
	4	146	93.0	158	8	US-10-713-381-3	Sequence 3, Appli
	5	50	31.8	50	8	US-10-713-381-5	Sequence 5, Appli
	6	40	25.5	40	8	US-10-713-381-6	Sequence 6, Appli
c	7	32.2	20.5	399	8	US-10-674-124A-9081	Sequence 9081, Ap
c	8	31.6	20.1	2352	7	US-10-282-122A-36272	Sequence 36272, A
	9	31.4	20.0	565	4	US-09-925-065A-871704	Sequence 871704,
c	10	30.8	19.6	601	4	US-09-925-065A-778732	Sequence 778732,
c	11	30.8	19.6	1122	7	US-10-437-963-1496	Sequence 1496, Ap
	12	30.6	19.5	559	5	US-10-027-632-244811	Sequence 244811,
	13	30.6	19.5	559	6	US-10-027-632-244811	Sequence 244811,
	14	30.6	19.5	79467	7	US-10-052-482-223	Sequence 223, App
	15	30.4	19.4	52001	7	US-10-317-273-11	Sequence 11, Appl
	16	30.2	19.2	436	4	US-09-925-065A-153052	Sequence 153052,
	17	30.2	19.2	1266	9	US-10-501-282-2923	Sequence 2923, Ap
	18	30.2	19.2	1754382	9	US-10-501-282-6651	Sequence 6651, Ap
	19	30	19.1	30	8	US-10-713-381-4	Sequence 4, Appli
	20	30	19.1	843	5	US-10-027-632-170626	Sequence 170626,
	21	30	19.1	843	6	US-10-027-632-170626	Sequence 170626,
	22	29.8	19.0	236	3	US-09-783-590-10490	Sequence 10490, A
	23	29.8	19.0	617	4	US-09-925-065A-285326	Sequence 285326,
	24	29.8	19.0	617	4	US-09-925-065A-285327	Sequence 285327,
	25	29.8	19.0	617	4	US-09-925-065A-285328	Sequence 285328,
c	26	29.8	19.0	691	7	US-10-424-599-57018	Sequence 57018, A
	27	29.8	19.0	1363	4	US-09-925-065A-12385	Sequence 12385, A
	28	29.8	19.0	2735	5	US-10-027-632-254612	Sequence 254612,
	29	29.8	19.0	2735	6	US-10-027-632-254612	Sequence 254612,
	30	29.8	19.0	1223197	5	US-10-027-632-179264	Sequence 179264,
	31	29.8	19.0	1223197	6	US-10-027-632-179264	Sequence 179264,
c	32	29.6	18.9	680	7	US-10-767-701-2154	Sequence 2154, Ap
c	33	29.2	18.6	1475	6	US-10-321-802-11	Sequence 11, Appl
c	34	29.2	18.6	1830121	7	US-10-329-670-1	Sequence 1, Appli
c	35	29.2	18.6	1830121	8	US-10-158-865-1	Sequence 1, Appli
c	36	29.2	18.6	1830121	9	US-10-981-687-1	Sequence 1, Appli
c	37	29	18.5	358	8	US-10-425-115-78606	Sequence 78606, A
	38	29	18.5	553	5	US-10-027-632-222444	Sequence 222444,
	39	29	18.5	553	5	US-10-027-632-222445	Sequence 222445,
	40	29	18.5	553	6	US-10-027-632-222444	Sequence 222444,
	41	29	18.5	553	6	US-10-027-632-222445	Sequence 222445,
c	42	29	18.5	7192	7	US-10-437-963-85365	Sequence 85365, A
	43	28.8	18.3	544	4	US-09-925-065A-805392	Sequence 805392,
	44	28.8	18.3	732	5	US-10-027-632-21721	Sequence 21721, A
	45	28.8	18.3	732	5	US-10-027-632-21722	Sequence 21722, A

ALIGNMENTS

RESULT 1
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%; Score 157; DB 8; Length 255;
Best Local Similarity 100.0%; Pred. No. 2.1e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 74

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 134

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db        135 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171
```

RESULT 2
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1

```
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1394
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-1

  Query Match            100.0%;  Score 157;  DB 8;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-43;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 3
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1394
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-2

  Query Match            100.0%;  Score 157;  DB 8;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-43;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCC 1214

Qy         61 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1215 AAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATG 1274

Qy        121 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db       1275 CAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 4
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 158
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-3

  Query Match            93.0%;  Score 146;  DB 8;  Length 158;
  Best Local Similarity   99.4%;  Pred. No. 1e-39;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 59
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy         60 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120

Qy        120 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              |||||||||||||||||||||||||||||||||||||
Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158


RESULT 5
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-5
```

```
  Query Match           31.8%;  Score 50;  DB 8;  Length 50;
  Best Local Similarity  100.0%;  Pred. No. 5.7e-07;
  Matches   50;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         75 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 50


RESULT 6
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIMOTHY W.
;  APPLICANT:  GARNAAT, CARL W.
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  KENDALL, TIMMY L.
;  TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
;  TITLE OF INVENTION:  USING SAME
;  FILE REFERENCE: 578R
;  CURRENT APPLICATION NUMBER: US/10/713,381
;  CURRENT FILING DATE:  2003-11-14
;  PRIOR APPLICATION NUMBER: 08/880,499
;  PRIOR FILING DATE: 1997-06-23
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Zea mays
US-10-713-381-6

  Query Match           25.5%;  Score 40;  DB 8;  Length 40;
  Best Local Similarity  100.0%;  Pred. No. 0.0014;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         85 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||||||||||||||||||||||||||||||||||||||
Db          1 AGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 40


RESULT 7
US-10-674-124A-9081/c
; Sequence 9081, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
;  APPLICANT: INOKO, Hidetoshi
;  APPLICANT:  TAMIYA, Gen
;  TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
;  TITLE OF INVENTION:  GENETIC POLYMORPHISM MARKERS
;  FILE REFERENCE: ORIN-003CIP
;  CURRENT APPLICATION NUMBER: US/10/674,124A
;  CURRENT FILING DATE:  2003-09-26
;  PRIOR APPLICATION NUMBER: 10/257,511
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: PCT/JP00/07621
;  PRIOR FILING DATE: 2000-10-30
;  PRIOR APPLICATION NUMBER: JP2000-112699
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: JP2002-327516
;  PRIOR FILING DATE: 2002-09-28
;  PRIOR APPLICATION NUMBER: JP2002-383869
;  PRIOR FILING DATE: 2002-12-09
;  NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 9081
;   LENGTH: 399
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: chr5.fa.07frz.139538206
;   FEATURE:
;   OTHER INFORMATION: Located on chromosome 5
;   FEATURE:
;   OTHER INFORMATION: Distance between a terminus base of telomere on
;   OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
;   OTHER INFORMATION: sequence : 128437212
;   FEATURE:
;   OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
;   OTHER INFORMATION: sequence listing upward to telomere on chrosomal short arm and
;   OTHER INFORMATION: 5'-terminus of this base sequence : 250186
US-10-674-124A-9081

  Query Match           20.5%;  Score 32.2;  DB 8;  Length 399;
  Best Local Similarity  61.2%;  Pred. No. 1.6;
  Matches   52;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              ||| ||||  ||  |    | || | |||| || |||  |||||||  |||||  |
Db        239 ATTTTGAAGAAAACAGTGGGCCACCCTCAAAAAACCCAAACTACTCATCTAACTTAAAAA 180

Qy        132 CAAACACGCACATATGTTTCCTGAA 156
              ||| | ||||| || | |  ||||
Db        179 AAAAAAACCACACATTTGTGATGAA 155


RESULT 8
US-10-282-122A-36272/c
; Sequence 36272, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36272
;   LENGTH: 2352
;   TYPE: DNA
;   ORGANISM: Streptococcus mutans
```

US-10-282-122A-36272

Query Match 20.1%; Score 31.6; DB 7; Length 2352;
Best Local Similarity 58.5%; Pred. No. 4.9;
Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
Qy         62 AGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGC 121
              | |||  |     |  || | | ||||  | ||||||  || ||||  |  |||   |||
Db        234 ATCCTGACTGTGCCCAAAGCCAAAGGAAGCATACTCCTGAAAAATCTGTGCTACATCTGC 175

Qy        122 AACTTCCATGCAAACACGCACATATGTTTCCTGA 155
              || |||     ||  |||  ||  | ||||||||
Db        174 AATTTCAGCAAAAGAACGTTCAATTTTTCCTGA 141
```

RESULT 9
US-09-925-065A-871704
; Sequence 871704, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 871704
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-871704

Query Match 20.0%; Score 31.4; DB 4; Length 565;
Best Local Similarity 60.2%; Pred. No. 3.4;
Matches 50; Conservative 1; Mismatches 32; Indels 0; Gaps 0;

```
Qy          4 GTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAG 63
              | | |||  | ||| | |   |||||||| |  | ||  | |||   ||| || |  ||
Db        274 GGCCTCTTGCTTGGAAGAGCCCATGCTTGGTTTAACGCCCTTCTGTCTCCGTCTTGAAAT 333

Qy         64 CCTTGCCTATTCTGAACCAAGAG 86
                | ||   ||| ||||| |:|||
Db        334 TCATGATAATTTTGAACAARGAG 356
```

RESULT 10
US-09-925-065A-778732/c
; Sequence 778732, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 778732
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-778732

Query Match 19.6%; Score 30.8; DB 4; Length 601;
Best Local Similarity 52.3%; Pred. No. 5.5;
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

```
Qy         12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
              |||| | | |||  ||    |  |  || ||| ||| | |    ||  |   | |||| |||
Db        354 ACATAGTACACTTTATTTATTGTTTACTGTTGGTCATTTAGGATATTTATACCCTTCCCT 295

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              ||| | ||  |  |   |  |||  | || ||| |   |||  |      | ||| | |
Db        294 ATTATAAATGATAAATTTTTCTATACACATACAGTTTTTCTACCCATATTATCTTTCCTA 235

Qy        132 CAAACACGCA 141
                | ||   ||
Db        234 AACACTGACA 225
```

RESULT 11
US-10-437-963-1496/c
; Sequence 1496, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 1496
; LENGTH: 1122
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101357C.1
US-10-437-963-1496

Query Match 19.6%; Score 30.8; DB 7; Length 1122;
Best Local Similarity 57.1%; Pred. No. 7;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy         60 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 119
              ||||||| || |  | || | |  | ||||  |||  | |||  || |||| ||  |  |
Db        313 CAAGCCTCGCATCATTTGCATCCCGGGGATCACTAAACTCAACAAAGCCATATTCATTCT 254

Qy        120 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 157
               ||  | |         ||| | |  || | | ||||||
Db        253 TCAGATCCACAAATCGCACCCTCCCATATCTGCTGAAC 216
```

RESULT 12
US-10-027-632-244811
; Sequence 244811, Application US/10027632

```
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244811
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-244811

  Query Match             19.5%;  Score 30.6;  DB 5;  Length 559;
  Best Local Similarity   51.9%;  Pred. No. 6.3;
  Matches   69;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy         24 ACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAA 83
              ||||||  |  |  |  ||||  ||    ||||  |  |  |    |  |  ||||     |
Db        380 ACATGCATTTCTAGCTTTTCGATTTTGCTCATCTTGCTGGAGGTACTTTTTCTAGATTCA 439

Qy         84 GAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACA 143
                | |  ||||||    |||  |  |||||    |  ||  |  |||  |||       ||||  |
Db        440 TCTGCTCCCTACCTTTAAAGATTCCATGGAAGTCCTACAAAATCCCATAGCACACTGAAC 499

Qy        144 TATGTTTCCTGAA 156
                |  |  |  |  ||
Db        500 ATTTTCTTCCCAA 512


RESULT 13
US-10-027-632-244811
; Sequence 244811, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 244811
;   LENGTH: 559
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-244811

  Query Match             19.5%;  Score 30.6;  DB 6;  Length 559;
  Best Local Similarity   51.9%;  Pred. No. 6.3;
  Matches   69;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy         24 ACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAA 83
              ||||||  |  |  |  ||||  ||    ||||  |  |  |    |  |  ||||     |
Db        380 ACATGCATTTCTAGCTTTTCGATTTTGCTCATCTTGCTGGAGGTACTTTTTCTAGATTCA 439

Qy         84 GAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACA 143
                | |  ||||||    |||  |  |||||    |  ||  |  |||  |||       ||||  |
Db        440 TCTGCTCCCTACCTTTAAAGATTCCATGGAAGTCCTACAAAATCCCATAGCACACTGAAC 499

Qy        144 TATGTTTCCTGAA 156
                |  |  |  |  ||
Db        500 ATTTTCTTCCCAA 512


RESULT 14
US-10-052-482-223
; Sequence 223, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
;  APPLICANT: Engelhard, Eric
;  APPLICANT:  Morris, David
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
;  FILE REFERENCE: A-71087/RMS/DCF
;  CURRENT APPLICATION NUMBER: US/10/052,482
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 241
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 223
;   LENGTH: 79467
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4099)..(4369)
;   OTHER INFORMATION: "n" at positions 4099 to 4369 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5502)..(5521)
;   OTHER INFORMATION: "n" at positions 5502 to 5521 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10089)..(10620)
;   OTHER INFORMATION: "n" at positions 10089 to 10620 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (13273)..(13370)
;   OTHER INFORMATION: "n" at positions 13273 to 13370 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (20762)..(20781)
;   OTHER INFORMATION: "n" at positions 20762 to 20781 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (30751)..(30916)
;   OTHER INFORMATION: "n" at positions 30751 to 30916 can be any base
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (46579)..(46772)
;   OTHER INFORMATION: "n" at positoins 46759 to 46772 can be any base
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (77382)..(77401)
;   OTHER INFORMATION: "n" at positions 77372 to 77401 can be any base
US-10-052-482-223

  Query Match             19.5%;  Score 30.6;  DB 7;  Length 79467;
  Best Local Similarity   56.4%;  Pred. No. 39;
  Matches   57;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy         51 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 110
              ||||     |  | ||||||  |  |||| | ||||  ||  | || || || |||  |
Db      53779 TCCACTAACTGATCCTTGCTTCCTCTGGAACAAGCTGAGTCTTAGTCACACACACACACA 53838

Qy        111 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTC 151
              |  || ||  ||    |||  || ||||   | | ||  ||
Db      53839 CACACACACACACACACCACACACACACAAGCTTGTGGCTC 53879


RESULT 15
US-10-317-273-11
; Sequence 11, Application US/10317273
; Publication No. US20040110158A1
; GENERAL INFORMATION:
;  APPLICANT: Kenneth W. Dobie
;  APPLICANT:  Tamara Balac Sipes
;  TITLE OF INVENTION: MODULATION OF MITOCHONDRIAL RIBOSOMAL PROTEIN L13 EXPRESSION
;  FILE REFERENCE: RTS-0478
;  CURRENT APPLICATION NUMBER: US/10/317,273
;  CURRENT FILING DATE:  2002-12-10
;  NUMBER OF SEQ ID NOS: 159
; SEQ ID NO 11
;   LENGTH: 52001
;   TYPE: DNA
;   ORGANISM: H. sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11121, 11122, 11123, 11124, 11125, 11126, 11127, 11128, 11129,
;   LOCATION: 11130, 11131, 11132, 11133, 11134, 11135, 11136, 11137,
;   LOCATION: 11138, 11139, 11140, 11141, 11142, 11143, 11144, 11145,
;   LOCATION: 11146, 11147, 11148, 11149, 11150, 11151, 11152, 11153
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11154, 11155, 11156, 11157, 11158, 11159, 11160, 11161, 11162,
;   LOCATION: 11163, 11164, 11165, 11166, 11167, 11168, 11169, 11170,
;   LOCATION: 11171, 11172, 11173, 11174, 11175, 11176, 11177, 11178,
;   LOCATION: 11179, 11180, 11181, 11182, 11183, 11184, 11185, 11186
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11187, 11188, 11189, 11190, 11191, 11192, 11193, 11194, 11195,
;   LOCATION: 11196, 11197, 11198, 11199, 11200, 11201, 11202, 11203,
;   LOCATION: 11204, 11205, 11206, 11207, 11208, 11209, 11210, 11211,
;   LOCATION: 11212, 11213, 11214, 11215, 11216, 11217, 11218, 11219
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 11220
;   OTHER INFORMATION: n = A,T,C or G
US-10-317-273-11

  Query Match             19.4%;  Score 30.4;  DB 7;  Length 52001;
  Best Local Similarity   61.2%;  Pred. No. 39;
  Matches   49;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy         50 TTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCA 109
              || | |  || ||  ||    |    ||||  ||||| ||| | | |   |||| | |||
Db      20055 TTTCTTTTTCTAAAACTGAAATTCCATGAAAAAAGAGAATATCAATTTTAAAACTACCCA 20114

Qy        110 TCTTACTCATGCAACTTCCA 129
              |||| ||||   |||||| |
Db      20115 TCTTTCTCAGCAAACTTCTA 20134
```

```
Search completed: March  6, 2006, 03:55:07
Job time : 227.751 secs
```

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 01:58:03 ; Search time 163.89 Seconds
(without alignments)
2100.307 Million cell updates/sec

Title: US-10-713-381-1_COPY_1155_1311
Perfect score: 157
Sequence: 1 cgtgtcatctcacatggcat.....cgcacatatgtttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	31.4	20.0	565	6	US-09-925-065A-871704	Sequence 871704,
c	2	30.8	19.6	601	6	US-09-925-065A-778732	Sequence 778732,
	3	30.6	19.5	64415	12	US-11-117-187-185	Sequence 185, App
c	4	30.4	19.4	191797	12	US-11-121-086-13	Sequence 13, Appl
	5	30.2	19.2	436	6	US-09-925-065A-153052	Sequence 153052,
	6	29.8	19.0	617	6	US-09-925-065A-285326	Sequence 285326,
	7	29.8	19.0	617	6	US-09-925-065A-285327	Sequence 285327,
	8	29.8	19.0	617	6	US-09-925-065A-285328	Sequence 285328,
	9	29.8	19.0	1363	6	US-09-925-065A-12385	Sequence 12385, A
c	10	29.4	18.7	611587	12	US-11-117-187-209	Sequence 209, App
	11	28.8	18.3	544	6	US-09-925-065A-805392	Sequence 805392,
	12	28.8	18.3	1370	8	US-10-750-185-47517	Sequence 47517, A
	13	28.8	18.3	1370	8	US-10-750-623-47517	Sequence 47517, A
	14	28.4	18.1	628	6	US-09-925-065A-845392	Sequence 845392,
c	15	28.4	18.1	1691	6	US-09-925-065A-33976	Sequence 33976, A
c	16	28.4	18.1	1691	6	US-09-925-065A-33977	Sequence 33977, A
c	17	28.4	18.1	1691	6	US-09-925-065A-33980	Sequence 33980, A
	18	28.2	18.0	2069	6	US-09-925-065A-724023	Sequence 724023,
	19	28	17.8	451	6	US-09-925-065A-173193	Sequence 173193,
c	20	28	17.8	801	6	US-09-925-065A-32050	Sequence 32050, A
c	21	28	17.8	801	6	US-09-925-065A-32051	Sequence 32051, A
c	22	28	17.8	1691	6	US-09-925-065A-33978	Sequence 33978, A
c	23	28	17.8	1691	6	US-09-925-065A-33979	Sequence 33979, A
c	24	28	17.8	1740	6	US-09-925-065A-553735	Sequence 553735,
c	25	28	17.8	1740	6	US-09-925-065A-553736	Sequence 553736,
c	26	28	17.8	2327	6	US-09-925-065A-686351	Sequence 686351,
c	27	28	17.8	2327	6	US-09-925-065A-686352	Sequence 686352,
c	28	28	17.8	2327	6	US-09-925-065A-686353	Sequence 686353,
c	29	28	17.8	2327	6	US-09-925-065A-686354	Sequence 686354,
c	30	28	17.8	2341	8	US-10-750-185-30272	Sequence 30272, A
c	31	28	17.8	2341	8	US-10-750-623-30272	Sequence 30272, A
	32	28	17.8	2750	8	US-10-750-185-39884	Sequence 39884, A
	33	28	17.8	2750	8	US-10-750-623-39884	Sequence 39884, A
	34	27.8	17.7	541	6	US-09-925-065A-427091	Sequence 427091,
c	35	27.8	17.7	560	6	US-09-925-065A-945788	Sequence 945788,
	36	27.8	17.7	576	6	US-09-925-065A-757513	Sequence 757513,
	37	27.8	17.7	1476	8	US-10-750-185-37809	Sequence 37809, A
	38	27.8	17.7	1476	8	US-10-750-623-37809	Sequence 37809, A
	39	27.8	17.7	3242	8	US-10-750-185-58624	Sequence 58624, A
	40	27.8	17.7	3242	8	US-10-750-623-58624	Sequence 58624, A
c	41	27.8	17.7	393323	7	US-10-330-773-23	Sequence 23, Appl
c	42	27.6	17.6	1724	8	US-10-750-185-55178	Sequence 55178, A
c	43	27.6	17.6	1724	8	US-10-750-623-55178	Sequence 55178, A
	44	27.6	17.6	2246	6	US-09-925-065A-68127	Sequence 68127, A
	45	27.4	17.5	531	6	US-09-925-065A-95915	Sequence 95915, A

ALIGNMENTS

RESULT 1
US-09-925-065A-871704
; Sequence 871704, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 871704
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-871704

Query Match 20.0%; Score 31.4; DB 6; Length 565;
Best Local Similarity 60.2%; Pred. No. 2.7;
Matches 50; Conservative 1; Mismatches 32; Indels 0; Gaps 0;

```
Qy          4 GTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAG 63
              | | |||  | ||| | |  |||||||| |  | ||  | |||   ||| || |  ||
Db        274 GGCCTCTTGCTTGGAAGAGCCCATGCTTGGTTTAACGCCCTTCTGTCTCCGTCTTGAAAT 333

Qy         64 CCTTGCCTATTCTGAACCAAGAG 86
                 | ||   ||| ||||| |:|||
Db        334 TCATGATAATTTTGAACAARGAG 356
```

RESULT 2

```
US-09-925-065A-778732/c
; Sequence 778732, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 778732
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-778732

  Query Match             19.6%;  Score 30.8;  DB 6;  Length 601;
  Best Local Similarity   52.3%;  Pred. No. 4.3;
  Matches   68;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy         12 ACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCT 71
              |||| | | |||  ||    | |  || ||| ||| | |   ||  |    | |||| |||
Db        354 ACATAGTACACTTTATTTATTGTTTACTGTTGGTCATTTAGGATATTTATACCCTTCCCT 295

Qy         72 ATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
              ||| | ||  |  |   |  |||  | || ||| |   |||  |     | ||| | |
Db        294 ATTATAAATGATAAATTTTTCTATACACATACAGTTTTTCTACCCATATTATCTTTCCTA 235

Qy        132 CAAACACGCA 141
              | ||   ||
Db        234 AACACTGACA 225


RESULT 3
US-11-117-187-185
; Sequence 185, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
;  APPLICANT: PREUSS, DAPHNE
;  APPLICANT:  COPENHAVER, GREGORY
;  TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
;  FILE REFERENCE: ARCD:309US
;  CURRENT APPLICATION NUMBER: US/11/117,187
;  CURRENT FILING DATE:  2005-04-28
;  PRIOR APPLICATION NUMBER: US/09/531,120
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/125,219
;  PRIOR FILING DATE: 1999-03-18
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 185
;   LENGTH: 64415
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (9960)..(21146)
;   OTHER INFORMATION: N = A, C G, or T/U
US-11-117-187-185

  Query Match             19.5%;  Score 30.6;  DB 12;  Length 64415;
  Best Local Similarity   58.1%;  Pred. No. 20;
  Matches   54;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy         65 CTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAAC 124
              ||||| | ||         | || |||  | |||| ||||| |  | || |  | | || |
Db       7117 CTTGCTTCTTTAAGTATTATAGTATATTTGCTCCTAAACACTAAACCTAAACCCTACACC 7176

Qy        125 TTCCATGCAAACACGCACATATGTTTCCTGAAC 157
              ||  || | ||  |  | || |  ||||| |||
Db       7177 TTAAATCCCAAACCCTAAATCTAATTCCTTAAC 7209


RESULT 4
US-11-121-086-13/c
; Sequence 13, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
;   LENGTH: 191797
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-13

  Query Match             19.4%;  Score 30.4;  DB 12;  Length 191797;
  Best Local Similarity   57.3%;  Pred. No. 32;
  Matches   55;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy         55 TCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTA 114
              || ||||   |  |||||| |||| |     |  | ||| | ||||||  |     ||
Db     162845 TCCTCCAGATCGCTCCTATTTTGAAATATATAAACGCTTACACACAAACACACTCATATA 162786

Qy        115 CTCATGCAACTTCCATGCAAACACGCACATATGTTT 150
              |||   ||| |||    | ||| ||||| | |||
Db     162785 TCCATTTTACTCCCACCACACCACACACATTTTTTT 162750


RESULT 5
US-09-925-065A-153052
; Sequence 153052, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 153052
;   LENGTH: 436
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

US-09-925-065A-153052

Query Match 19.2%; Score 30.2; DB 6; Length 436;
Best Local Similarity 60.2%; Pred. No. 6.1;
Matches 50; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 4 GTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAG 63
Db 256 GGCCTCTTGCTTGGAAGAGCCCATGCTTGGTTTAACGCCCTTCTGTCTCCGTCTTGAAAT 315

Qy 64 CCTTGCCTATTCTGAACCAAGAG 86
Db 316 TCATGATAATTTTGAACAAGGAG 338

RESULT 6
US-09-925-065A-285326
; Sequence 285326, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 285326
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-285326

Query Match 19.0%; Score 29.8; DB 6; Length 617;
Best Local Similarity 60.5%; Pred. No. 9;
Matches 49; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 74 TCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCA 133
Db 373 TGTGTAGAAATTTAATACATACTTCCAAACAATCCAGAGTTTTGTAGCAACTTATACTCC 432

Qy 134 AACACGCACATATGTTTCCTG 154
Db 433 TACCTGCACTRAAGTTAAATG 453

RESULT 7
US-09-925-065A-285327
; Sequence 285327, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 285327
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-285327

Query Match 19.0%; Score 29.8; DB 6; Length 617;
Best Local Similarity 60.5%; Pred. No. 9;
Matches 49; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 74 TCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCA 133
Db 373 TGTGTAGAAATTTAATACATACTTCCAAACAATCCAGAGTTTTGTAGCAACTTATACTCC 432

Qy 134 AACACGCACATATGTTTCCTG 154
Db 433 TACCTGCACTGAAGTTAAATG 453

RESULT 8
US-09-925-065A-285328
; Sequence 285328, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 285328
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-285328

Query Match 19.0%; Score 29.8; DB 6; Length 617;
Best Local Similarity 60.5%; Pred. No. 9;
Matches 49; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 74 TCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCA 133
Db 373 TGTGTAGAAATTTAATACATACTTCCAAACAATCCAGAGTTTTGTAGCAACTTATACTCC 432

Qy 134 AACACGCACATATGTTTCCTG 154
Db 433 TACCTGCACTGAAGTTAAATG 453

RESULT 9
US-09-925-065A-12385
; Sequence 12385, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:

```
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12385
;   LENGTH: 1363
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-12385

  Query Match             19.0%;  Score 29.8;  DB 6;  Length 1363;
  Best Local Similarity   50.3%;  Pred. No. 11;
  Matches   73;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          6 CATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCC 65
              || | ||| ||||  | ||   | ||  | |||| ||      |||   || || |
Db        765 CACCGCACCTGGCCAAGTATCAGTTTTCTTAACCATTGTCACTGTCTTATGGTTCTTCTG 824

Qy         66 TTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACT 125
              |    | |||||   || ||| ||     |      |||  |||     |   | |||||  |
Db        825 TGTTGTCTTCTGTGGCATGAGTATGTACAGGAAAAAAAGATCGGCTTATATAATGCATTT 884

Qy        126 TCCATGCAAACACGCACATATGTTT 150
                |||  || |  ||  || || |
Db        885 GAGATGGGAAAATTCAAGTACGTGT 909


RESULT 10
US-11-117-187-209/c
; Sequence 209, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
;  APPLICANT: PREUSS, DAPHNE
;  APPLICANT:  COPENHAVER, GREGORY
;  TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
;  FILE REFERENCE: ARCD:309US
;  CURRENT APPLICATION NUMBER: US/11/117,187
;  CURRENT FILING DATE:  2005-04-28
;  PRIOR APPLICATION NUMBER: US/09/531,120
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/125,219
;  PRIOR FILING DATE: 1999-03-18
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 209
;   LENGTH: 611587
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-11-117-187-209

  Query Match             18.7%;  Score 29.4;  DB 12;  Length 611587;
  Best Local Similarity   52.0%;  Pred. No. 91;
  Matches   66;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy          8 TCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTT 67
              |||| | || ||  | |   | ||  |||  | || ||      |   || | |  ||
Db     270144 TCTCTCCTGCCAAGCGAATGGATTTGTCATGCATTTGTTGCTAACTGCCGACACAAGCTG 270085

Qy         68 GCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTTC 127
                || | ||  |||  ||     |  || || |  |  |||    |||||||   | ||| |
Db     270084 GCTTGTTAGGAAAAAAACCCTTTTCTTCTTCTTATGAATGTTGCTTACTCTGTCCACTGC 270025

Qy        128 CATGCAA 134
                || |||
Db     270024 GATACAA 270018


RESULT 11
US-09-925-065A-805392
; Sequence 805392, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 805392
;   LENGTH: 544
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-805392

  Query Match             18.3%;  Score 28.8;  DB 6;  Length 544;
  Best Local Similarity   69.6%;  Pred. No. 18;
  Matches   39;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         71 TATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTT 126
              ||||  |||  ||||  | ||| |   ||||| ||||||  ||| |  ||||||  ||
Db         23 TATTAAGAAAAAAGAAAAAACCAAAAACCAAAAAATCCAACTTTCCAATGCAAATT 78


RESULT 12
US-10-750-185-47517
; Sequence 47517, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47517
;   LENGTH: 1370
;   TYPE: DNA
;   ORGANISM: Bovine  19866880570548
US-10-750-185-47517

  Query Match             18.3%;  Score 28.8;  DB 8;  Length 1370;
```

```
  Best Local Similarity   54.8%;  Pred. No. 24;
  Matches   57;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          47 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 106
               ||||| | || | |  |||   || |||| |   | ||| | |    ||||  |||  ||
Db         660 TTGTTTCCTCATGCTTGCCGCTCCCATTCAGGGACTAGAAGCTGAAAACTCTGAAATTAT 719

Qy         107 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTT 150
               || | |    |  | || ||     || | | ||| |   ||||
Db         720 CCTTGTGCAACTAGGAAATTGTGGACAGAAATGCAAAATAGTTT 763


RESULT 13
US-10-750-623-47517
; Sequence 47517, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47517
;   LENGTH: 1370
;   TYPE: DNA
;   ORGANISM: Bovine  19866880570548
US-10-750-623-47517

  Query Match             18.3%;  Score 28.8;  DB 8;  Length 1370;
  Best Local Similarity   54.8%;  Pred. No. 24;
  Matches   57;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          47 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 106
               ||||| | || | |  |||   || |||| |   | ||| | |    ||||  |||  ||
Db         660 TTGTTTCCTCATGCTTGCCGCTCCCATTCAGGGACTAGAAGCTGAAAACTCTGAAATTAT 719

Qy         107 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTT 150
               || | |    |  | || ||     || | | ||| |   ||||
Db         720 CCTTGTGCAACTAGGAAATTGTGGACAGAAATGCAAAATAGTTT 763


RESULT 14
US-09-065A-845392
; Sequence 845392, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 845392
;   LENGTH: 628
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-845392

  Query Match             18.1%;  Score 28.4;  DB 6;  Length 628;
  Best Local Similarity   54.9%;  Pred. No. 25;
  Matches   56;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy          30 TTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGAT 89
               | || |   | | |  | ||    || |  ||| ||||   |||   ||    | || |
Db         495 TGGTGCTGGCCTCCAGATAGTGAGTTCTTGTAAGACTTGGTCATTTAAAAGTGTGCGGCT 554

Qy          90 ACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
               || ||| ||  ||  ||  |||| ||| |||   | |||||
Db         555 CCCCACTGCCCCACTCTCTCTCTTGCTCCTGCCTTTGCCATG 596


RESULT 15
US-09-925-065A-33976/c
; Sequence 33976, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33976
;   LENGTH: 1691
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-33976

  Query Match             18.1%;  Score 28.4;  DB 6;  Length 1691;
  Best Local Similarity   54.9%;  Pred. No. 34;
  Matches   56;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy          30 TTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGAT 89
               | || |   | | |  | ||    || |  ||| ||||   |||   ||    | || |
Db        1191 TGGTGCTGGCCTCCAGATAGTGAGTTCTTGTAAGACTTGGTCATTTAAAAGTGTGCGGCT 1132

Qy          90 ACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATG 131
               || ||| ||  ||  ||  |||| ||| |||   | |||||
Db        1131 CCCCACTGCCCCACTCTCTCTCTTGCTCCTGCCTTTGCCATG 1090


Search completed: March  6, 2006, 10:32:44
Job time : 164.89 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 17:38:03 ; Search time 5721.35 Seconds
(without alignments)
13025.213 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311
Perfect score: 1311
Sequence: 1 ccatggtgtctctatgaaaa.....cgcacatatgtttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1311	100.0	1394	6	BD062176	BD062176 Male tiss
2	1311	100.0	1394	6	BD062177	BD062177 Male tiss
3	1311	100.0	1394	6	AX224394	AX224394 Sequence
4	1311	100.0	1394	6	AX224395	AX224395 Sequence
5	1311	100.0	3343	15	AF360356	AF360356 Zea mays
6	160.6	12.3	255	6	AX224402	AX224402 Sequence
7	146	11.1	158	6	AX224396	AX224396 Sequence
8	77.2	5.9	188283	14	AC155584	AC155584 Zea mays
9	77	5.9	186199	14	AC147602	AC147602 Zea mays
10	74.4	5.7	161616	14	CR936840	CR936840 Danio rer
11	73.4	5.6	25769	2	AC117267	AC117267 Dictyoste
c 12	72.4	5.5	172293	14	AC155379	AC155379 Zea mays
c 13	71.6	5.5	201985	14	AC155618	AC155618 Zea mays
c 14	71.2	5.4	713	2	AF034389	AF034389 Plasmodiu
c 15	71.2	5.4	3576	2	DDU86962	U86962 Dictyosteli
16	71	5.4	317137	15	AY664413	AY664413 Zea mays
c 17	70.8	5.4	149598	5	BX957346	BX957346 Zebrafish
c 18	70.8	5.4	156452	14	AC155376	AC155376 Zea mays
c 19	70.4	5.4	347582	2	PFMAL4P1	AL034557 Plasmodiu
c 20	70.2	5.4	883	6	CQ422641	CQ422641 Sequence
c 21	70	5.3	241	10	BV119878	BV119878 PZA01377
c 22	70	5.3	247	10	BV119874	BV119874 PZA01377
c 23	70	5.3	248	10	BV119877	BV119877 PZA01377
24	69.8	5.3	13400	6	CQ855052	CQ855052 Sequence
c 25	69.6	5.3	134692	14	AC151050	AC151050 Zea mays
c 26	69.4	5.3	110000	15	CR382132_17	Continuation (18 o
27	69.2	5.3	112468	14	AC149836	AC149836 Zea mays
28	69.2	5.3	266544	2	AC116956	AC116956 Dictyoste
c 29	69	5.3	148120	14	AC155517	AC155517 Zea mays
c 30	68.8	5.2	7347	1	AF211124	AF211124 Carsonell
c 31	68.6	5.2	186199	14	AC147602	AC147602 Zea mays
c 32	68.4	5.2	213	10	BV119875	BV119875 PZA01377
c 33	68.4	5.2	216	10	BV119871	BV119871 PZA01377
c 34	68.4	5.2	224	10	BV119882	BV119882 PZA01377
c 35	68.4	5.2	245	10	BV119881	BV119881 PZA01377
c 36	68.4	5.2	246	10	BV119869	BV119869 PZA01377
37	68.2	5.2	346296	15	AF090447	AF090447 Zea mays
38	68	5.2	1092	6	CS123057	CS123057 Sequence
39	68	5.2	110000	14	PFMAL7P1_02	Continuation (3 of
40	67.8	5.2	227495	14	AC148480	AC148480 Zea mays
c 41	67.6	5.2	351	10	BV110530	BV110530 PZA01059
42	67.6	5.2	110000	15	AY664419_2	Continuation (3 of
43	67.6	5.2	331039	2	AC116979	AC116979 Dictyoste
c 44	67.4	5.1	170	10	BV119873	BV119873 PZA01377
c 45	67.2	5.1	155263	14	BX927284	BX927284 Danio rer

ALIGNMENTS

RESULT 1
BD062176
LOCUS BD062176 1394 bp DNA linear PAT 27-AUG-2002
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062176
VERSION BD062176.1 GI:22607781
KEYWORDS JP 2001520523-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC ,C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES Location/Qualifiers
source 1. .1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 7.2e-228;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

RESULT 2
BD062177

```
LOCUS       BD062177               1394 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Male tissue-preferred regulatory region and method of using same.
ACCESSION   BD062177
VERSION     BD062177.1  GI:22607782
KEYWORDS    JP 2001520523-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1394)
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
            Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: JP 2001520523-A 2 30-OCT-2001;
            PIONEER HI BRED INTERNATIONAL INC
COMMENT     PN   JP 2001520523-A/2
            PD   30-OCT-2001
            PF   19-JUN-1998 JP 1999504910
            PR   23-JUN-1997 US    08/880499
            PI   MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
            PI   TIMMY L KENDALL
            PC   C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
            ,C07K14/34,C12Q1/68,
            PC   A01H5/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES                Location/Qualifiers
     source             1. .1394
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN
```

Query Match 100.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 7.2e-228;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
```

```
Qy      241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy      301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy      361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy      421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy      481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy      541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy      601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy      661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy      721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy      781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy      841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy      901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy      961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy     1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy     1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy     1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy     1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy     1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
```

```
RESULT 3
AX224394
LOCUS       AX224394                1394 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 1 from Patent WO0160997.
ACCESSION   AX224394
VERSION     AX224394.1  GI:15554636
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 1 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1394
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             100.0%;  Score 1311;  DB 6;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-228;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTAGTAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTAGTAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
```

```
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Db        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311

Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 4
AX224395
LOCUS       AX224395               1394 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 2 from Patent WO0160997.
ACCESSION   AX224395
VERSION     AX224395.1  GI:15554637
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 2 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1394
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             100.0%;  Score 1311;  DB 6;  Length 1394;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-228;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Db        781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Db        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
```

```
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 5
AF360356
LOCUS       AF360356               3343 bp    DNA     linear   PLN 12-MAY-2001
DEFINITION  Zea mays male fertility protein (Ms45) gene, complete cds.
ACCESSION   AF360356
VERSION     AF360356.1  GI:14028756
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 3343)
  AUTHORS   Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
  TITLE     Cloning of Ms45, a gene required for male ferility from Zea mays
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3343)
  AUTHORS   Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
            Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
            IA 50131-1004, USA
FEATURES             Location/Qualifiers
     source          1. .3343
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /chromosome="9L"
     gene            <1392. .>2942
                     /gene="Ms45"
     mRNA            join(<1392. .1768,1898. .2182,2280. .2447,2534. .>2942)
                     /gene="Ms45"
                     /product="male fertility protein"
     CDS             join(1392. .1768,1898. .2182,2280. .2447,2534. .2942)
                     /gene="Ms45"
                     /codon_start=1
                     /product="male fertility protein"
                     /protein_id="AAK52489.1"
                     /db_xref="GI:14028757"
                     /translation="MEKRNLQWRRGRDGIVQYPHLFFAALALALLVADPFGLSPLAEV
                     DYRPVKHELAPYGEVMGSWPRDNASRLRRGRLEFVGEVFGPESIEFDLQGRGPYAGLA
                     DGRVVRWMGEEAGWETFAVMNPDWSEEVCANGVNSTTRKQHEKEEFCGRPLGLRFHGE
                     TGELYVADAYYGLMVVGQSGGVASSVAREADGDPIRFANDLDVHRNGSVFFTDTSMRY
                     SRKDHLNILLEGEGTGRLLRYDPETSGVHVVLKGLVFPNGVQISEDHQFLLFSETTNC
                     RIMRYWLEGPRAGEVEVFANLPGFPDNVRSNGRGQFWVAIDCCRTPAQEVFAKRPWLR
                     TLYFKFPLSLKVLTWKAARRMHTVLALLDGEGRVVEVLEDRGHEVMKLVSEVREVGRK
                     LWIGTVAHNHIATIPYPLED"
ORIGIN

  Query Match             100.0%;  Score 1311;  DB 15;  Length 3343;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-228;
  Matches 1311;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGTGTCTCTATGAAAAAGATGAGTACAATGTGTCTATATCCGTTTTCTTAGGGTCC 60

Qy         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCTTCTGCCTTATTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTGCATGAT 120

Qy        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCATGTTCCACTTCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCAT 180

Qy        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCA 240

Qy        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCCACAATCGTCACCTATA 300

Qy        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATTATTCTTTAGATATTATTTAATTTTTGGAAAAATAACAAACTTATACTTTTGTGTA 360

Qy        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCCAGAAATGCGAGGACCAGCCATGTCTAGTG 420

Qy        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACT 480

Qy        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAA 540

Qy        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAA 600

Qy        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTT 660

Qy        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA 720

Qy        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC 780

Qy        781 TTTCTGATTTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTCTGATTTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTT 840

Qy        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Qy        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960

Qy        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTAT 1020

Qy       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCC 1080

Qy       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCA 1140

Qy       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC 1200

Qy       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAAT 1260

Qy       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311


RESULT 6
AX224402
LOCUS       AX224402                 255 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 9 from Patent WO0160997.
ACCESSION   AX224402
VERSION     AX224402.1  GI:15554644
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 9 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .255
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match            12.3%;  Score 160.6;  DB 6;  Length 255;
  Best Local Similarity  97.6%;  Pred. No. 4.5e-19;
  Matches  163;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1145 CGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGT 1204
                | || | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 CCGCGGATCCCGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGT 64

Qy       1205 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 1264
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 TCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCAT 124

Qy       1265 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 CTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 171


RESULT 7
AX224396
LOCUS       AX224396                 158 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 3 from Patent WO0160997.
ACCESSION   AX224396
VERSION     AX224396.1  GI:15554638
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
  TITLE     Male tissue-preferred regulatory region and method of using same
  JOURNAL   Patent: WO 0160997-A 3 23-AUG-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .158
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match            11.1%;  Score 146;  DB 6;  Length 158;
  Best Local Similarity  99.4%;  Pred. No. 2.2e-16;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1155 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTC-TTGTTCCATCGTC 1213
                ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db          1 CGTGTCATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTTGTTCCATCGTC 60

Qy       1214 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 1273
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGCCTTGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCAT 120

Qy       1274 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 1311
                |||||||||||||||||||||||||||||||||||||
Db        121 GCAACTTCCATGCAAACACGCACATATGTTTCCTGAAC 158


RESULT 8
AC155584
LOCUS       AC155584               188283 bp    DNA     linear   HTG 25-JAN-2005
DEFINITION  Zea mays strain B73 clone ZMMBBc0196I14, *** SEQUENCING IN PROGRESS
            ***, 20 unordered pieces.
ACCESSION   AC155584
VERSION     AC155584.2  GI:58082443
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 188283)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Consortium for Maize Genomics - BAC skim sequencing and assembly
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 188283)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
REFERENCE   3  (bases 1 to 188283)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
COMMENT     On Jan 25, 2005 this sequence version replaced gi:57863105.
            -------------- Trace submission
                Center_name: TIGR
                Seq_lib_id: ZGFX
            -------------- Project informaton
                Web site: http://www.tigr.org/tdb/tgi/maize/
                Contact: maize@tigr.org
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 20 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
```

```
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    22133: contig of 22133 bp in length
            *    22134    22233: gap of unknown length
            *    22234    27964: contig of 5731 bp in length
            *    27965    28064: gap of unknown length
            *    28065    35160: contig of 7096 bp in length
            *    35161    35260: gap of unknown length
            *    35261    49754: contig of 14494 bp in length
            *    49755    49854: gap of unknown length
            *    49855    96501: contig of 46647 bp in length
            *    96502    96601: gap of unknown length
            *    96602    98675: contig of 2074 bp in length
            *    98676    98775: gap of unknown length
            *    98776   103999: contig of 5224 bp in length
            *   104000   104099: gap of unknown length
            *   104100   123477: contig of 19378 bp in length
            *   123478   123577: gap of unknown length
            *   123578   126464: contig of 2887 bp in length
            *   126465   126564: gap of unknown length
            *   126565   134044: contig of 7480 bp in length
            *   134045   134144: gap of unknown length
            *   134145   142804: contig of 8660 bp in length
            *   142805   142904: gap of unknown length
            *   142905   154520: contig of 11616 bp in length
            *   154521   154620: gap of unknown length
            *   154621   159716: contig of 5096 bp in length
            *   159717   159816: gap of unknown length
            *   159817   160976: contig of 1160 bp in length
            *   160977   161076: gap of unknown length
            *   161077   162085: contig of 1009 bp in length
            *   162086   162185: gap of unknown length
            *   162186   167405: contig of 5220 bp in length
            *   167406   167505: gap of unknown length
            *   167506   171505: contig of 4000 bp in length
            *   171506   171605: gap of unknown length
            *   171606   173244: contig of 1639 bp in length
            *   173245   173344: gap of unknown length
            *   173345   184423: contig of 11079 bp in length
            *   184424   184523: gap of unknown length
            *   184524   188283: contig of 3760 bp in length.
FEATURES             Location/Qualifiers
     source          1. .188283
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBc0196I14"
     gap             22134. .22233
                     /estimated_length=unknown
     gap             27965. .28064
                     /estimated_length=unknown
     gap             35161. .35260
                     /estimated_length=unknown
     gap             49755. .49854
                     /estimated_length=unknown
     gap             96502. .96601
                     /estimated_length=unknown
     gap             98676. .98775
                     /estimated_length=unknown
     gap             104000. .104099
                     /estimated_length=unknown
     gap             123478. .123577
                     /estimated_length=unknown
     gap             126465. .126564
                     /estimated_length=unknown
     gap             134045. .134144
                     /estimated_length=unknown
     gap             142805. .142904
                     /estimated_length=unknown
     gap             154521. .154620
                     /estimated_length=unknown
     gap             159717. .159816
                     /estimated_length=unknown
     gap             160977. .161076
                     /estimated_length=unknown
     gap             162086. .162185
                     /estimated_length=unknown
     gap             167406. .167505
                     /estimated_length=unknown
     gap             171506. .171605
                     /estimated_length=unknown
     gap             173245. .173344
                     /estimated_length=unknown
     gap             184424. .184523
                     /estimated_length=unknown
ORIGIN

  Query Match             5.9%;  Score 77.2;  DB 14;  Length 188283;
  Best Local Similarity   80.5%;  Pred. No. 0.00021;
  Matches  103;  Conservative    0;  Mismatches   23;  Indels    2;  Gaps    1;

Qy        789 TTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAA 848
              || ||||||| |||||||||    ||||||||  ||| || | ||||| ||||| | ||| |
Db      35512 TTGTTTAAGGGCTAGTTTAAGGACCCTGTT--TTTAAAGGGATTTTTATTTTCTTAAGAG 35569

Qy        849 AAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGAC 908
              |||||||||||||||| || |  |||||||||| | |||||||||||| |||| ||| ||
Db      35570 AAATTAGTTTATTTTTCATAGTAAAATAGAAATCTCTTAGAAAAATATAGTTCCCAAAC 35629

Qy        909 TAGCCCTA 916
              ||||||||
Db      35630 TAGCCCTA 35637


RESULT 9
AC147602
LOCUS       AC147602               186199 bp    DNA     linear   HTG 17-AUG-2004
DEFINITION  Zea mays clone ZMMBBc0334A01, *** SEQUENCING IN PROGRESS ***, 6
            ordered pieces.
ACCESSION   AC147602
VERSION     AC147602.5  GI:51315585
KEYWORDS    HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 186199)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Butler,E., Wing,R., Bharti,A.K.
            and Messing,J.
  TITLE     Zea mays, clone ZMMBBc0334A01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 186199)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (19-DEC-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 186199)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Butler,E., Wing,R., Bharti,A.K.,
            Messing,J., Abouelleil,A., Allen,N., Anderson,M., Anderson,S.,
            Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S.,
            Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P.,
            FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L.,
            Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P.,
            Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-AUG-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Aug 17, 2004 this sequence version replaced gi:49658659.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@broad.mit.edu
            --------------
                Bharti, AK and Messing, J: The Plant Genome Initiative at
            Rutgers, Waksman Institute, Rutgers, The State University of New
            Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            (http://pgir.rutgers.edu)
                Butler, E and Wing, R: Arizona Genomics Institute, Biological
            Sciences West, 448A, P.O. Box 210088, University of Arizona,
            Tucson, AZ 85721, USA (http://www.genome.arizona.edu)
            -------------- Project Information
                Center project name: L30003
                Center clone name: 334_A_1
            -------------- Consensus Information
            This consensus is derived from a shotgun assembly that has been
            manually curated. It is the best representation of the BAC that we
            can generate without further laboratory work. The draft assembly
            has been edited, and if possible, ends identified by vector as well
            as by BAC end sequences, and contigs ordered and oriented. Bases
            that are not Ns are either above Q20 or manually edited. This
            assembly was performed with Arachne (Genome Res. 2002 12: 177-189;
            Genome Res. 2003 13: 91-96). All trace files for this project are
            available at the NCBI trace repository
            (http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?). An exact list of
            reads used in this assembly are available at
            http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   100617: contig of 100617 bp in length
            *  100618   100717: gap of unknown length
            *  100718   104730: contig of 4013 bp in length
            *  104731   104830: gap of unknown length
            *  104831   115104: contig of 10274 bp in length
            *  115105   115204: gap of unknown length
            *  115205   156396: contig of 41192 bp in length
            *  156397   156496: gap of unknown length
            *  156497   179936: contig of 23440 bp in length
            *  179937   180036: gap of unknown length
            *  180037   186199: contig of 6163 bp in length.
FEATURES             Location/Qualifiers
     source          1. .186199
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBc0334A01"
                     /clone_lib="CHORI-201 Maize B73 BAC MboI Library"
     gap             100618. .100717
                     /estimated_length=unknown
     gap             104731. .104830
                     /estimated_length=unknown
     gap             115105. .115204
                     /estimated_length=unknown
     gap             156397. .156496
                     /estimated_length=unknown
     gap             179937. .180036
                     /estimated_length=unknown
ORIGIN

  Query Match             5.9%;  Score 77;  DB 14;  Length 186199;
  Best Local Similarity   71.4%;  Pred. No. 0.00023;
  Matches  115;  Conservative    0;  Mismatches   45;  Indels    1;  Gaps    1;

Qy        780 CTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTT 839
              ||||     ||  ||| || |||||||||| |||||| ||| ||  ||  |||||| ||||
Db     162058 CTTTTCTTTTACTTTTAGGGCTAGTTTGGGAACCCTATTT-TTCTAAGAAATTTTCATTT 162116

Qy        840 TTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAG 899
              ||  ||   ||||||||| |||||||||| | ||||||||||    || | |||||||||
Db     162117 TTCTAAGTGAAATTAGTTCATTTTCTCTTAACAAAATAGAAATTTATTGGTAAAATAGAG 162176

Qy        900 TTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAA 940
              || ||| |||| ||||  | | | |    |||  |  ||||
Db     162177 TTTCCAAACTAACCCTCAAGTATATAATGAATGCAGGACAA 162217


RESULT 10
CR936840
LOCUS       CR936840              161616 bp    DNA     linear   HTG 26-FEB-2005
DEFINITION  Danio rerio clone DKEY-91O21, *** SEQUENCING IN PROGRESS ***, 13
            unordered pieces.
ACCESSION   CR936840
VERSION     CR936840.2  GI:60302473
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 161616)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 26, 2005 this sequence version replaced gi:60279457.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zK91O21
            -------------- Summary Statistics
```

```
        Assembly program: XGAP4; version 4.5
        Chemistry: Dye-terminator; 100% of reads
        Consensus quality: 156361 bases at least Q40
        Consensus quality: 157330 bases at least Q30
        Consensus quality: 158179 bases at least Q20
        Insert size: 160416; sum-of-contigs
        Insert size: 160313; 4.7% error; agarose-fp
        Quality coverage: 6.55x in Q20 bases; sum-of-contigs Quality
        coverage: 6.56x in Q20 bases; agarose-fp
        ----------.
        * NOTE: This is a 'working draft' sequence. It currently
        * consists of 13 contigs. The true order of the pieces
        * is not known and their order in this sequence record is
        * arbitrary. Gaps between the contigs are represented as
        * runs of N, but the exact sizes of the gaps are unknown.
        * This record will be updated with the finished sequence
        * as soon as it is available and the accession number will
        * be preserved.
        *        1     3158: contig of 3158 bp in length
        *     3159     3258: gap of 100 bp
        *     3259    11860: contig of 8602 bp in length
        *    11861    11960: gap of 100 bp
        *    11961    21800: contig of 9840 bp in length
        *    21801    21900: gap of 100 bp
        *    21901    45034: contig of 23134 bp in length
        *    45035    45134: gap of 100 bp
        *    45135    47371: contig of 2237 bp in length
        *    47372    47471: gap of 100 bp
        *    47472    66727: contig of 19256 bp in length
        *    66728    66827: gap of 100 bp
        *    66828    79067: contig of 12240 bp in length
        *    79068    79167: gap of 100 bp
        *    79168    97393: contig of 18226 bp in length
        *    97394    97493: gap of 100 bp
        *    97494   103542: contig of 6049 bp in length
        *   103543   103642: gap of 100 bp
        *   103643   112675: contig of 9033 bp in length
        *   112676   112775: gap of 100 bp
        *   112776   137298: contig of 24523 bp in length
        *   137299   137398: gap of 100 bp
        *   137399   158790: contig of 21392 bp in length
        *   158791   158890: gap of 100 bp
        *   158891   161616: contig of 2726 bp in length.
FEATURES             Location/Qualifiers
     source          1. .161616
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-91O21"
                     /clone_lib="DanioKey"
     misc_feature    1. .3158
                     /note="assembly_fragment:00020
                     fragment_chain:1"
     misc_feature    3259. .11860
                     /note="assembly_fragment:00240
                     fragment_chain:1"
     misc_feature    11961. .21800
                     /note="assembly_fragment:00342
                     fragment_chain:1"
     misc_feature    21901. .45034
                     /note="assembly_fragment:01299
                     fragment_chain:1"
     misc_feature    45135. .47371
                     /note="assembly_fragment:00002
                     fragment_chain:1"
     misc_feature    47472. .66727
                     /note="assembly_fragment:00606
                     fragment_chain:1"
     misc_feature    66828. .79067
                     /note="assembly_fragment:00463
                     fragment_chain:1"
     misc_feature    79168. .97393
                     /note="assembly_fragment:00794
                     fragment_chain:1"
     misc_feature    97494. .103542
                     /note="assembly_fragment:00074
                     fragment_chain:1"
     misc_feature    103643. .112675
                     /note="assembly_fragment:00146
                     fragment_chain:1"
     misc_feature    112776. .137298
                     /note="assembly_fragment:01585
                     fragment_chain:1"
     misc_feature    137399. .158790
                     /note="assembly_fragment:01032
                     fragment_chain:1"
     misc_feature    158891. .161616
                     /note="assembly_fragment:00044.0"
ORIGIN

  Query Match             5.7%;  Score 74.4;  DB 14;  Length 161616;
  Best Local Similarity   47.9%;  Pred. No. 0.00071;
  Matches  184;  Conservative    0;  Mismatches  200;  Indels    0;  Gaps    0;

Qy        576 TTTACAAACATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTT 635
              ||   |||     || |   | |||  |     ||  |     |     | | |  |
Db      79191 TTATAAAATNNNNAANAAAAAAAAAAAAAATATATAAAATTTTTNNTANNTTNNATTTAAT 79250

Qy        636 TCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAG 695
              | |||||  || || ||  |||| || |   ||||    ||   | |.||   ||  ||||
Db      79251 TATTTTTTTTTTTTTTTTTTTTTTTATATTTTTATTATTATATTTTTATTTTTAATAT 79310

Qy        696 AGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCA 755
              | ||||| ||    |||  |||| | | |||  | |  ||||| | |          |
Db      79311 TTTTATTTTTTTATAATTAAATAATAAAAATATTTATTTTTTTTTTTTTTTTTTTAAAA 79370

Qy        756 TGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCT 815
              |   |  | |  |  |||  | |||  ||  |||| | |  ||  |    |    |
Db      79371 ATTAAAATTATTTTNTNAAAATTTTTTTTTNTTTATTTTTATATATATATATATATATAT 79430

Qy        816 GTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAA 875
              | | ||| ||| ||  ||     |||     |   |||  ||| |||| | ||| |   |
Db      79431 TTATATTTAAAAAAAAATTTTANTTTTTTTTATTTATTTTTTTTTTTTTTTTTTTNTA 79490

Qy        876 TAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAAT 935
              |  |||| |  || | | ||| |  |     ||   || |   | |    ||| ||
Db      79491 TNAAAAAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 79550

Qy        936 TACAATCACTGTGTATAATTATTT 959
                ||   | | ||  ||| |||
Db      79551 ATATATATATATATAATATTTTTT 79574


RESULT 11
AC117267
LOCUS       AC117267               25769 bp    DNA     linear   INV 21-FEB-2004
DEFINITION  Dictyostelium discoideum chromosome 2 map 5836255-5862024 strain
            AX4, complete sequence.
ACCESSION   AC117267 AC115597
VERSION     AC117267.2  GI:42733680
KEYWORDS    HTG.
SOURCE      Dictyostelium discoideum
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 25769)
  AUTHORS   Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P.,
            Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K.,
            Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
            Noegel,A.A.
  TITLE     Sequence and analysis of chromosome 2 of Dictyostelium discoideum
  JOURNAL   Nature 418 (6893), 79-85 (2002)
   PUBMED   12097910
  REMARK    The Dictyostelium Genome Sequencing Consortium
REFERENCE   2  (bases 1 to 25769)
  AUTHORS   Baumgart,C.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (09-APR-2002) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
REFERENCE   3  (bases 1 to 25769)
  AUTHORS   Baumgart,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2004) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
COMMENT     On or before Feb 21, 2004 this sequence version replaced
            gi:19570016, gi:20087114.
            CDS predictions from GeneID may contain errors. Further Information
            is available from IMB Jena, Department of Genome Analysis
            (http://genome.imb-jena.de/dictyostelium/)
            and the Univerity Colonge, Institute for Biochemistry I
            (http://www.uni-koeln.de/dictyostelium/project.shtml
            Funding
            Agency : Deutsche Forschungsgemeinschaft (DFG).
FEATURES             Location/Qualifiers
     source          1. .25769
                     /organism="Dictyostelium discoideum"
                     /mol_type="genomic DNA"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /chromosome="2"
                     /map="5836255-5862024"
     CDS             complement(join(832. .1070,1237. .1349,1533. .1798,
                     1896. .2053,2231. .2360))
                     /note="GeneID exon scores (in order of location ranges):
                     23.68, 1.54, 15.75, 9.21, -4.75 - GSCJ_ID dd_03387"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAS38625.1"
                     /db_xref="GI:42733681"
                     /translation="MRLQIILFLSILLIFQLKSTFSSNDISSTILSITPNFNKDYNNQ
                     ILSKSEWFRSLEYSSNKKVIFEVFDGFNHEYIENLDRIIKEFRENKVYGVDFSFSFCL
                     NSKFNYLEVIKKFSTISRNNKVYPENIWLNIESTHENCYNNSNYLYNKNNNEYQEYLV
                     FSKQVIKSFLQNGYNIGIKSSKNTFIKIFENNKDKLFSSLPFFYLKNIAENQEFEDYY
                     SSPLGSWTDFKIKQFKKLNDIETILSDVEVNLCSSSNNNNNNNNVDDETLTLLGSGSG
                     SGSGSGSSGTSGAQLSGTGYVINHQ"
     CDS             join(2957. .3173,3272. .5583)
                     /note="GeneID exon scores (in order of location ranges):
                     24.88, 254.96 - GSCJ_ID dd_00753"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAS38626.1"
                     /db_xref="GI:42733682"
                     /translation="MEFPTKYVYITSNLDKPKEWFTQSEMSLTTTTDTIQKSFVNNSG
                     SSSSSKGFGEAVLLDILDTAGQEEYSAMRDQYVRTGDCFMIVFSIDSRSSFEEVSQLK
                     QHIERVKDRDDVPIIIVGNKVDLESRRQVSRIEADQLARSLRVPYIETSAKTRSNIEE
                     AFFTLVRHTPRNTVYKVVVMGGGGVGKSAIIIQFIQNHFVEEYDPTIEDSYRKQVTIS
                     GLPPIGGSLNKKSSSSSSSSSSSSSGKTGLFNKIFSGKDKQPSPQQAASPSTIDRTGQI
                     STNRLEANVLSYSMSNLSKEVPLITGDCVYCQGCNVILSRFSNLVKTGDDSFTWKCEF
                     CKYSNSNILLEQGEIPNKDSVEYVLSSPSTSSTTDGSKREESIIIYCIDVSGSMGITT
                     EVPSLQSEWVNAKKGVKGASSGPSYISRLECVQSSIPTMIDRLSIQYPNKRVVLVTFS
                     DEVMIYTQSNSVDGPIVIAGDKLEDFDQLIEIGRSMTYDKLPTASGSSDFLKAKIKSL
                     EPVQSTALGPALLVSAAIASQKMLSEVVICTDGVPNVGLGAIEDLPLGPAQEFYEKVT
                     KLAQNNKTTINIIGIGSGSHIDLGVIGKVSEQTNGNITIIHPLELAREIRKLTQNPMIA
                     TDVEMSICLHPTLEINKYDSKQGLSRVVKQFPNVNSLTDLTLLYSSRNRPTEFVQIYP
                     FQIQIKYTKLDGVRCLRVVSAQLQATPDFNTSTSNANISILAMAFTQQAAKIAQQQEY
                     MESRLHLKAATKLIRSLCNTDEQWEEFYNFEVLREEMEAPLITCIKNKQQRVEKAATD
                     DEIQVFYKMKNVHKSFVEGGRKKDISRRKGEAEINKQYYNIKFT"
     CDS             complement(5711. .6796)
                     /note="GeneID exon scores (in order of location ranges):
                     68.61 - GSCJ_ID dd_00752"
                     /codon_start=1
                     /product="similar to Plasmodium falciparum (isolate 3D7).
                     Hypothetical protein"
                     /protein_id="AAS38627.1"
                     /db_xref="GI:42733683"
                     /translation="MSSRVKSLIERYNGINLTIEKSLSPRGLINTNINNNINNNNNNN
                     GTITTTTSGGNNNNNSTSPSTNPVSFDRVVDHSNNNQIQIVIDSFVNINNNNNNNNNN
                     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNFNINQNKTNNQNQNQNHLNCN
                     KFNNNNNYYFDFSTNEMITTLNSCTLEQIGTIGNDCYSTNTTPSPTSSSSSTSSLLNN
                     KNSNHWMMPSLKINQIINDNNNNISNINFVQIQSNSSSLSSSCSTYQSCPIEDEHIQE
                     QLRIQKEQLKQIDDDDDDDDDDDDDDDSSPSPPPPPTTTTKTFCLSPRPIIKKEENID
                     QNSMSPPPKLDKRRFSRDLIPFTIANI"
     CDS             join(8924. .9076,9236. .10573)
                     /note="GeneID exon scores (in order of location ranges):
                     10.37, 128.60 - GSCJ_ID dd_00729"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAS38628.1"
                     /db_xref="GI:42733684"
                     /translation="MSSSTIITPIATTTTNNNSSTNGIHNSKTNRTLFKNIKNNIPNS
                     PNNAPIKKNSVGIIGNGPNVNNNNKDRKLNSNDGVNFGKKNILFTPSKNSMVSTTKHS
                     SSSSSLLAKNEYDIESIKKELKSMDISTGGISPLSSPLRESSPISTIDLESITATTTK
                     LTDAPVAASSSSSSVKYTIDVLLSFKSANTKRPIQIDIVENHQKEIVVPLSLPTTPYN
                     NNNNNNNNTNNSSQHNTANGQKYPIFSPQISPFKIAYAQSSSTKTNINNNNNNNNITP
                     TKKANSNITTPQSNNRNNHYNSNTKSSTKKQNPIPFSLNTATPNQKNTTTTPSKKSTT
                     TTPKSNKKINDVNAAFAAVANSTTTTEQPVAIIDGATPSTNESPKKTPLKETDPKRLA
                     ARQRQIDIGKMTAGYKNYIALVPKSKRKPTDPKTPNKNQVCSKRSWDGQIKKWRRQLH
                     ENYDWTGIPFNEKEVEASNNESTEESTLEKEMNKLEISEENENHNN"
     CDS             complement(10972. .11979)
                     /note="GeneID exon scores (in order of location ranges):
                     23.29 - GSCJ_ID dd_00728"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAS38629.1"
                     /db_xref="GI:42733685"
                     /translation="MEEKIKFYFEIIDFQNQKFKIQEFTSKLIGLKEESFTTFKPIVY
                     EKYLNWTQSIEESILKTNGTINKSIFEEIFSYCGYIGEFLEYEPFIVFLKFTTFGIIL
                     DDYIFEKINSLNMKLNEKEKLINSLIYYNNKNENKIGFEEFWEIINEFQNFTHKESFQR
                     IINANSLWIKSSIDSRNISNSPINSRCSFNEYFEKRSSDASGDFILTISMIGIMDNYI
                     ENSIIESKEFQTINYHAKSFFLLINDLYSFNREINENDLLNYIKILAIQLNSIQLSID
                     KTIELIIDHYLKFLSSIETILKLYQNDQSTYQLLKQVFQNSNKILSGIYFAHKKSKRY
                     N"
     CDS             join(12884. .13020,13172. .15320)
                     /note="GeneID exon scores (in order of location ranges):
                     2.42, 136.56 - GSCJ_ID dd_00727"
                     /codon_start=1
                     /product="similar to Dictyostelium discoideum (Slime
                     mold). Hypothetical 97.7 kDa protein"
                     /protein_id="AAS38630.1"
                     /db_xref="GI:42733686"
                     /translation="MDNLIYDFLIKKEYNQQSFLDCILILSINGLYISYDGLLDYCNN
                     FRLIENKTVNKEAISYYGLNYNQYQLLDEFTRLLKLKNKTTDCKNKIINTHNSLNKCC
                     GNNKFDYSILIIIENNNNIINKDYQFNNNNKNSNNNIYNNNNDNNNSIIEEKDLNKII
                     TLKNERINGNDGEITTTTTTTTTSRSSTTIMTTTTTTNGVNDRNYLLFRQVWRNIVIK
                     TQILFHLRLYNIHANKKVFLTPIQLVDYKFKGYLQSMVLNYHDDNDDDADEDSYDSDI
                     DDDDSYSDGGNCSSSSSGSSDIGSSSNSINNGISNSSSSSILSNSSLLELSMIRVNSL
                     PEGLKSIEFEKEYNVIQDRLLPPSISSIRFSYGFNQRIAKGVISDNVISITFGDSFNQ
                     SLDGNWLPKQLKHLQFGHKFQQTIKMGQLPSSLTSLILDPRSYKGVIEIGSIPDSVKT
                     LDYKFNSCSNQESISFNPIPKSITRLVFDSEFNQNIKANDISSNYNLTSIHFGEHFNS
                     DIGIKSLPNSIREIKFGRAFDRDIKLCPSSITSIDFGNKFNRPLSMMTQTLTSIDFGS
                     KFNQIIPQGIFIHTKLKSLNFGYHFNQIIPADTLPPTLESLNLGGYNREITVKNDEYD
                     CYGISNKGGFGSNSSSNFCVGGTNNGLREMLKNTTSLKTLTLNYFNRKIEVGDLPNSI
                     ESLNLGYHFNQPIGNNVLPKLLKKLFILNSEFNQNISADGCIPFGLQTIYIRNSNMNF
                     INSLNPLFITKYINIIDLSHL"
     CDS             join(16184. .16199,17272. .17613,17707. .19331)
                     /note="GeneID exon scores (in order of location ranges):
                     0.08, 32.32, 227.42 - GSCJ_ID dd_00726"
                     /codon_start=1
                     /product="similar to Dictyostelium discoideum (Slime
                     mold). Spore coat protein SP87 precursor (PL3 protein)"
                     /protein_id="AAS38631.1"
                     /db_xref="GI:42733687"
                     /translation="MFIGEDCNKITNEEECHKSSECIVINYTPCCGEQKWACSKGTFD
                     TCTYENSCYRNSSNNQVVEVSNKCFNLDGFIKITTPTEYSCSDAKIKECALLGKSCSF
                     QKNSCSNPTSCCPGESICEGLSSGSSTSGGGSSGGTSGGSSSGGTSGGSSSGGTSGSS
                     SSGSSSGGVSSCSTTHCPEGYHCSMVNDVATCLASTTGGTGLPGTSSSTAGVSSCLTT
                     LCPIGHICVEDSNGVNCVPNGGGTSGGSSSTGTSGGHPDPCRDVDCPDGFHCECKDGK
                     TAKCVPSPTTGSSSTSGGHPDPCKDVTCPDGFHCECKDGKTAKCVPSPTTGSSSTSGG
                     NTNPCSNVNCPDGFYCECKDGKTAKCVPSGPTQPPKPPVCSLRCPPNHECRFNDQGHQ
                     CCVKVHHDRCSLRCPHGHECKVDHNGKECCVRSHRPPPPEVCSLRCPPKHECKFDDHG
                     KKCCVKIHCDEVCDLDCGRGFECKIRHDGSKCCVRSERPHPPQHEKCNKRCPPGHECK
                     VDQHGKECCVVAHRPPPKCSLRCPPRHECRVNHFGEECCVKVHHDKCSLRCPPGHECK
                     VDQHGKECCVVAHRPPPKCSLRCPPKHECRINHFGEECCVKSRNDCLTCEDLNCERKG
```

```
                LHCAMKTVPIDKENCCEKVPVCYSNNPLLDGGHGFI"
     CDS        19762. .21336
                /note="GeneID exon scores (in order of location ranges):
                106.68 - GSCJ_ID dd_00725"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAS38632.1"
                /db_xref="GI:42733688"
                /translation="MDNNNLPQPYFPQSNNYCDCDDNENGNDSTLTIQPKPIAPSLPS
                SQQPSPTIISQINHNKYQLFKEFKEHNNNIDSKMLCVQLSEFQISPKTATNTFDSWLN
                SLWLPPTGFKKLVDNHLKKVYIPFYSYRVNYNVEFSGFVTYDSKLNTNALYDKYGNTN
                STYNNNNYYNVNNNSLNNSNNADKNKKTITSTSATILNKENVELYNNGNNNNLNNKNS
                EDILKDKSIKWNKVDGYITRTINRILLLACNSMDLLSDESIMNCMVSQQFDFNNVIEK
                SLDEFNQFAQLDQVEVLARYEPDVVWKRDYLNHLLETEQSMSIEYLKQQEQSKFVKNV
                SSTIQFNEKLFSLVYIPFYYTKYEFQNQSFSFFLNSQNGVISATRPEIGLGKIGDFLK
                YTQNYFSTLVGNFEPPFSKCKGDELAIIDSFKVYNTSSYFLTFSRSSTNFLGSSSIGHL
                TIRNNSKSITATLRGQKRRSVQKGNPFSLLPLHERTFSFKGHWCFEIIDGEVNQIDVI
                KINPREGGNIYQSIKN"
     CDS        complement(21961. .24357)
                /note="GeneID exon scores (in order of location ranges):
                173.48 - GSCJ_ID dd_00724"
                /codon_start=1
                /product="similar to Arabidopsis thaliana (Mouse-ear
                cress). Putative PREG1-like negative regulator"
                /protein_id="AAS38633.1"
                /db_xref="GI:42733689"
                /translation="MSKVYSKSQTVSPGKPQNIFVTALPSSVSAPSLLSLQKPNNGAG
                SHRKQRQNSGSSKNRPNLPLDQIKSTERKNSGNKSNNNNNKNRKYSKNNSNVNSNNIN
                SSSIVIKCEEQQQQQQQQQQQQQQQQQQQQQQQQQQDGFLKSQEFPNLFKNKLYLKDDIS
                SPSSSSSTSSPSCLSSSSSSTFSQSEIAENKDDINNNNCNNTNGNFEEEEEEEEKNKV
                EQQSIEEETFIESSFEETKEVVEEVPFEEQFEILPISIGEETLKEYTSFLSVLSELIV
                DMYNDDDKSSTSSQSSSSSSSSSAPSSPKPIKNNDNNTTTTTTTATTTTSKTATSS
                LSTIEKFQVFNVDQIPEISIEAYIERVFKYLPFGTDIFIFSTIYLDRLIQWNQEIQIS

  Query Match          5.6%;  Score 73.4;  DB 2;  Length 25769;
  Best Local Similarity  49.4%;  Pred. No. 0.0014;
  Matches  217;  Conservative    0;  Mismatches  221;  Indels    1;  Gaps    1;

Qy        438 CCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGAAAGCTATCATGTAA 497
              | | |  | || | |||||| ||     ||     |   |    |||   | |   ||
Db       2392 CGAAACAACGACTAAAAAAAAAAAAAATTTGTTGAGGAAAAAAAAAAAAAAAAAAAAAAA 2451

Qy        498 TGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTTCGAAAG 557
              | |  ||| ||| |   || ||||  | | |  |   |||||| ||  |||| |  |||
Db       2452 TATGAAAAAAAAAAAAAATAAAAAAAAAGGTTTTTACAAAAAAAAAAAATATCTAAAAAA 2511

Qy        558 AGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAAGGTAAG 617
              | | | ||||   |   |||| |   || |||  |  |  ||| | |    || |||
Db       2512 AAATTATTTATTAAATTATTTAAATTGATTTAATTGTTAACAATAAAAAAACAAAATAAA 2571

Qy        618 CTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACAT 677
              ||  |  || |  |  || ||    ||| || |  | || ||||| ||  |  | | ||
Db       2572 ATAAAATATTAATTTAATTATTAAA-ATTATTTTCTTATTTTTATTTTTAATTGAAAAAT 2630

Qy        678 TTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATT 737
              |   || |  | |     ||  |||| |||  |||| |   |  | |  | |  | ||
Db       2631 TAATTTTTAAAAGATATTTGTTTTTTTTTTCTTTTTTTTTTTTTTTTAAAGTGATCTATTT 2690

Qy        738 TTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAG 797
              || | |  | |        | |  ||  || || ||| | ||   || | | |
Db       2691 TTAAAAATTGGGTGTTTGATTGAAAAATATTCTTTTAAATTTTTTCAGGAAATATCTTTT 2750

Qy        798 AGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTT 857
                | |||   ||| |      |||       ||||  |||||| ||||||  | | ||
Db       2751 TTTTTTTTTCAAACACCACCAATTTTTTTTTTTTTTTTTTTTTTAAAAAACCTAAATT 2810

Qy        858 TATTTTCTCTTTATAAAAT 876
              |  ||| | ||  |   ||
Db       2811 TTATTTTTTTATTTTTAT 2829


RESULT 12
AC155379/c
```

```
LOCUS       AC155379             172293 bp    DNA     linear   HTG 25-JAN-2005
DEFINITION  Zea mays strain B73 clone ZMMBBb0131C15, *** SEQUENCING IN PROGRESS
            ***, 27 unordered pieces.
ACCESSION   AC155379
VERSION     AC155379.2  GI:58082241
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 172293)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Consortium for Maize Genomics - BAC skim sequencing and assembly
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 172293)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
REFERENCE   3  (bases 1 to 172293)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
COMMENT     On Jan 25, 2005 this sequence version replaced gi:57862900.
            -------------- Trace submission
                Center_name: TIGR
                Seq_lib_id: ZOCF
            -------------- Project informaton
                Web site: http://www.tigr.org/tdb/tgi/maize/
                Contact: maize@tigr.org
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 27 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2421: contig of 2421 bp in length
            *     2422     2521: gap of unknown length
            *     2522     5334: contig of 2813 bp in length
            *     5335     5434: gap of unknown length
            *     5435     8098: contig of 2664 bp in length
            *     8099     8198: gap of unknown length
            *     8199    10614: contig of 2416 bp in length
            *    10615    10714: gap of unknown length
            *    10715    12841: contig of 2127 bp in length
            *    12842    12941: gap of unknown length
            *    12942    15185: contig of 2244 bp in length
            *    15186    15285: gap of unknown length
            *    15286    36327: contig of 21042 bp in length
            *    36328    36427: gap of unknown length
            *    36428    44529: contig of 8102 bp in length
            *    44530    44629: gap of unknown length
            *    44630    58058: contig of 13429 bp in length
            *    58059    58158: gap of unknown length
            *    58159    66276: contig of 8118 bp in length
            *    66277    66376: gap of unknown length
```

```
          *    66377    70448: contig of 4072 bp in length
          *    70449    70548: gap of unknown length
          *    70549    74167: contig of 3619 bp in length
          *    74168    74267: gap of unknown length
          *    74268    82175: contig of 7908 bp in length
          *    82176    82275: gap of unknown length
          *    82276    83496: contig of 1221 bp in length
          *    83497    83596: gap of unknown length
          *    83597   114544: contig of 30948 bp in length
          *   114545   114644: gap of unknown length
          *   114645   116257: contig of 1613 bp in length
          *   116258   116357: gap of unknown length
          *   116358   119322: contig of 2965 bp in length
          *   119323   119422: gap of unknown length
          *   119423   121368: contig of 1946 bp in length
          *   121369   121468: gap of unknown length
          *   121469   125359: contig of 3891 bp in length
          *   125360   125459: gap of unknown length
          *   125460   126905: contig of 1446 bp in length
          *   126906   127005: gap of unknown length
          *   127006   131580: contig of 4575 bp in length
          *   131581   131680: gap of unknown length
          *   131681   134564: contig of 2884 bp in length
          *   134565   134664: gap of unknown length
          *   134665   138587: contig of 3923 bp in length
          *   138588   138687: gap of unknown length
          *   138688   156638: contig of 17951 bp in length
          *   156639   156738: gap of unknown length
          *   156739   167111: contig of 10373 bp in length
          *   167112   167211: gap of unknown length
          *   167212   168221: contig of 1010 bp in length
          *   168222   168321: gap of unknown length
          *   168322   172293: contig of 3972 bp in length.
FEATURES             Location/Qualifiers
     source          1. .172293
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0131C15"
     gap             2422. .2521
                     /estimated_length=unknown
     gap             5335. .5434
                     /estimated_length=unknown
     gap             8099. .8198
                     /estimated_length=unknown
     gap             10615. .10714
                     /estimated_length=unknown
     gap             12842. .12941
                     /estimated_length=unknown
     gap             15186. .15285
                     /estimated_length=unknown
     gap             36328. .36427
                     /estimated_length=unknown
     gap             44530. .44629
                     /estimated_length=unknown
     gap             58059. .58158
                     /estimated_length=unknown
     gap             66277. .66376
                     /estimated_length=unknown
     gap             70449. .70548
                     /estimated_length=unknown
     gap             74168. .74267
                     /estimated_length=unknown
     gap             82176. .82275
                     /estimated_length=unknown
     gap             83497. .83596
                     /estimated_length=unknown
     gap             114545. .114644
                     /estimated_length=unknown
     gap             116258. .116357
                     /estimated_length=unknown
     gap             119323. .119422
                     /estimated_length=unknown
     gap             121369. .121468
                     /estimated_length=unknown
     gap             125360. .125459
                     /estimated_length=unknown
     gap             126906. .127005
                     /estimated_length=unknown
     gap             131581. .131680
                     /estimated_length=unknown
     gap             134565. .134664
                     /estimated_length=unknown
     gap             138588. .138687
                     /estimated_length=unknown
     gap             156639. .156738
                     /estimated_length=unknown
     gap             167112. .167211
                     /estimated_length=unknown
     gap             168222. .168321
                     /estimated_length=unknown
ORIGIN

  Query Match             5.5%;  Score 72.4;  DB 14;  Length 172293;
  Best Local Similarity   78.6%;  Pred. No. 0.0016;
  Matches   99;  Conservative    0;  Mismatches   26;  Indels    1;  Gaps    1;

Qy        791 TTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAA 850
              | ||||||||||||||||||||| | |||| || ||| ||||||  |||||  |||  |||
Db      27185 TGTTAAGAGCTAGTTTGGCAACTCCGTTT-TTCCAAGGAATTTCTATTTTCCCAAGGAAA 27127

Qy        851 ATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTA 910
              ||||||| |||| |||||    |||||    || |   |||||||| ||||| ||| ||||
Db      27126 ATTAGTTCATTTACTCTTGGGAAAATGAGAATCCTCTAGAAAAATGGAGTTTCCAAACTA 27067

Qy        911 GCCCTA 916
              ||||||
Db      27066 GCCCTA 27061


RESULT 13
AC155618/c
LOCUS       AC155618              201985 bp    DNA     linear   HTG 25-JAN-2005
DEFINITION  Zea mays strain B73 clone ZMMBBc0344D04, *** SEQUENCING IN PROGRESS
            ***, 21 unordered pieces.
ACCESSION   AC155618
VERSION     AC155618.2  GI:58082477
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 201985)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Consortium for Maize Genomics - BAC skim sequencing and assembly
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 201985)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
REFERENCE   3  (bases 1 to 201985)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
```

```
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
COMMENT     On Jan 25, 2005 this sequence version replaced gi:57863139.
            ------------- Trace submission
                Center_name: TIGR
                Seq_lib_id: ZGGW
            ------------- Project informaton
                Web site: http://www.tigr.org/tdb/tgi/maize/
                Contact: maize@tigr.org
            -------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 21 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1916: contig of 1916 bp in length
            *    1917     2016: gap of unknown length
            *    2017    22480: contig of 20464 bp in length
            *   22481    22580: gap of unknown length
            *   22581    30468: contig of 7888 bp in length
            *   30469    30568: gap of unknown length
            *   30569    65218: contig of 34650 bp in length
            *   65219    65318: gap of unknown length
            *   65319    78774: contig of 13456 bp in length
            *   78775    78874: gap of unknown length
            *   78875    87797: contig of 8923 bp in length
            *   87798    87897: gap of unknown length
            *   87898    94463: contig of 6566 bp in length
            *   94464    94563: gap of unknown length
            *   94564   121761: contig of 27198 bp in length
            *  121762   121861: gap of unknown length
            *  121862   124636: contig of 2775 bp in length
            *  124637   124736: gap of unknown length
            *  124737   131029: contig of 6293 bp in length
            *  131030   131129: gap of unknown length
            *  131130   138739: contig of 7610 bp in length
            *  138740   138839: gap of unknown length
            *  138840   146859: contig of 8020 bp in length
            *  146860   146959: gap of unknown length
            *  146960   152969: contig of 6010 bp in length
            *  152970   153069: gap of unknown length
            *  153070   155026: contig of 1957 bp in length
            *  155027   155126: gap of unknown length
            *  155127   156434: contig of 1308 bp in length
            *  156435   156534: gap of unknown length
            *  156535   173224: contig of 16690 bp in length
            *  173225   173324: gap of unknown length
            *  173325   174858: contig of 1534 bp in length
            *  174859   174958: gap of unknown length
            *  174959   182700: contig of 7742 bp in length
            *  182701   182800: gap of unknown length
            *  182801   185872: contig of 3072 bp in length
            *  185873   185972: gap of unknown length
            *  185973   199501: contig of 13529 bp in length
            *  199502   199601: gap of unknown length
            *  199602   201985: contig of 2384 bp in length.
FEATURES             Location/Qualifiers
     source          1. .201985
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBc0344D04"
     gap             1917. .2016
                     /estimated_length=unknown
     gap             22481. .22580
                     /estimated_length=unknown
     gap             30469. .30568
                     /estimated_length=unknown
     gap             65219. .65318
                     /estimated_length=unknown
     gap             78775. .78874
                     /estimated_length=unknown
     gap             87798. .87897
                     /estimated_length=unknown
     gap             94464. .94563
                     /estimated_length=unknown
     gap             121762. .121861
                     /estimated_length=unknown
     gap             124637. .124736
                     /estimated_length=unknown
     gap             131030. .131129
                     /estimated_length=unknown
     gap             138740. .138839
                     /estimated_length=unknown
     gap             146860. .146959
                     /estimated_length=unknown
     gap             152970. .153069
                     /estimated_length=unknown
     gap             155027. .155126
                     /estimated_length=unknown
     gap             156435. .156534
                     /estimated_length=unknown
     gap             173225. .173324
                     /estimated_length=unknown
     gap             174859. .174958
                     /estimated_length=unknown
     gap             182701. .182800
                     /estimated_length=unknown
     gap             185873. .185972
                     /estimated_length=unknown
     gap             199502. .199601
                     /estimated_length=unknown
ORIGIN

  Query Match             5.5%;  Score 71.6;  DB 14;  Length 201985;
  Best Local Similarity   76.9%;  Pred. No. 0.0022;
  Matches  100;  Conservative    0;  Mismatches   29;  Indels    1;  Gaps    1;

Qy        792 TTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAA 851
              | |||| |||||||||| |||| | ||| || ||| | ||||| |||||  |||| |||||
Db      44986 TATAAGGGCTAGTTTGGGAACCATATTT-TTCCAAGGGATTTTCATTTTCCCAAAGAAAA 44928

Qy        852 TTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAG 911
              |||||| |||||| | |   |||||||||| | | | | |||||  |||| ||| |||||
Db      44927 TTAGTTCATTTTCCCATAGGAAAATAGAAATCCCATGGGAAAATGTAGTTCCCAAACTAG 44868

Qy        912 CCCTAGAATG 921
              ||||| || |
Db      44867 CCCTAAAAAG 44858


RESULT 14
AF034389/c
LOCUS       AF034389                 713 bp    DNA     linear   INV 04-FEB-1999
DEFINITION  Plasmodium falciparum sexual stage antigen (s16) gene, promoter and
            partial cds.
ACCESSION   AF034389
VERSION     AF034389.1  GI:3098290
KEYWORDS    .
SOURCE      Plasmodium falciparum (malaria parasite P. falciparum)
  ORGANISM  Plasmodium falciparum
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 713)
  AUTHORS   Dechering,K.J., Kaan,A.M., Mbacham,W., Wirth,D.F., Eling,W.,
            Konings,R.N. and Stunnenberg,H.G.
  TITLE     Isolation and functional characterization of two distinct
            sexual-stage-specific promoters of the human malaria parasite
            Plasmodium falciparum
  JOURNAL   Mol. Cell. Biol. 19 (2), 967-978 (1999)
   PUBMED   9891033
```

```
REFERENCE   2  (bases 1 to 713)
  AUTHORS   Dechering,K.J., Kaan,A.M. and Konings,R.N.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-1997) Molecular Biology, University of Nijmegen,
            Toernooiveld 1, Nijmegen 6525 ED, The Netherlands
FEATURES             Location/Qualifiers
     source          1. .713
                     /organism="Plasmodium falciparum"
                     /mol_type="genomic DNA"
                     /strain="NF54"
                     /db_xref="taxon:5833"
     gene            1. .>713
                     /gene="s16"
     promoter        1. .536
                     /gene="s16"
     mRNA            537. .>713
                     /gene="s16"
     5'UTR           537. .710
                     /gene="s16"
     CDS             711. .>713
                     /gene="s16"
                     /note="Pfs16"
                     /codon_start=1
                     /product="sexual stage antigen"
                     /protein_id="AAD12581.1"
                     /db_xref="GI:3098291"
                     /translation="M"
ORIGIN

  Query Match             5.4%;  Score 71.2;  DB 2;  Length 713;
  Best Local Similarity   45.8%;  Pred. No. 0.0063;
  Matches  282;  Conservative    0;  Mismatches  333;  Indels    1;  Gaps    1;

Qy        444 CAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGAAAGCTATCATGTAATGTTTA 503
              ||| |    |||||||||  ||   ||  |||  | |  |||    |  ||  |||||
Db        681 CAAAAGAAAAAAAAATAAAAAATAAAAAAAAATAAAATAAAAAAAAAAAAATAAAATGTAAG 622

Qy        504 AAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTTCGAAAGAGACAA 563
               || |     |||| | ||  | | |   | |  |||     |    |||    ||    | |
Db        621 AATATTATCCTATGATAATAAAGCTATTGTGGAAATTATATTTATGGAAAATATAATAT 562

Qy        564 ATTATGTTACAGTTACAAACATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATGA 623
               ||      || |  ||         | | | |   | ||  |||||  | |  ||   | |
Db        561 ATATATGTAAATATTTTTTTTTTATTATATATA-AATATATATATATATATATATTTTTT 503

Qy        624 CGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTT 683
                ||    |||||  ||||  ||   | ||||||| || || ||||| ||||||| || |
Db        502 TTTTTTTTTTTTTTATTTTATTTATTTATTTATTTATTTTTTTTTTTTTATAAATATAAAA 443

Qy        684 CTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATA 743
                | |  | |||  ||| |   |||          | | | || |||  |  |||   ||  |||
Db        442 TTTTCAGTATAAAAAGTATTTAACGGAAATATGTAGAATGGAAAAGAAGAATTTGGTATG 383

Qy        744 TAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAG 803
               |||          ||    || |     |     | ||| ||| |      ||   | |   ||
Db        382 TAAATTTTAATATAAAATATAAAAAAGAAATATAATTTTTTTATAAATTAATATAAAAAG 323

Qy        804 TTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTT 863
               |||     |     |  |    |  ||     ||||  |  |||| ||    | ||   |||| |||
Db        322 TTTATTTATTTTTATATATATAATTTTTTTTTAAATTTAAATTCATATATTTTTAATTT 263

Qy        864 CTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTT 923
                | ||| |      |               |||| |  | |||     |      |      |    | ||
Db        262 ATATTTTTTTTTTTTTTTTTTGCTTATATATATAATATATATACGTAATATAATATAATTT 203

Qy        924 TTCCCAATAAATTACAATCACTGTGTATAATTATTTGGCCAGCCCCATAAATTATTTAAA 983
               ||     |     ||   |            |||| ||||||      |     |||| || | | | |
Db        202 TTTTTTTTTTTTTTTTAATTTGAAATATATTTATTTTATTTTCTACATATATAAATAATA 143

Qy        984 CCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGA 1043
                 || | | ||      |          ||     | | ||| |    |     |   ||||| |
Db        142 TAAAATTTTATATTTTAAGGTGTATCATTAATGTTTTTTTAAAATAAAAAAAAAAATAGT 83

Qy       1044 GAGAGAGAGGAAGAAA 1059
                 |   |  |    | |||
Db         82 TATCTATATATAAAAA 67


RESULT 15
DDU86962/c
LOCUS       DDU86962                3576 bp    DNA     linear   INV 23-JUL-2001
DEFINITION  Dictyostelium discoideum MigA (migA) gene, complete cds.
ACCESSION   U86962
VERSION     U86962.1  GI:1841871
KEYWORDS    .
SOURCE      Dictyostelium discoideum
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 3576)
  AUTHORS   Escalante,R., Wessels,D., Soll,D.R. and Loomis,W.F.
  TITLE     Chemotaxis to cAMP and slug migration in Dictyostelium both depend
            on migA, a BTB protein
  JOURNAL   Mol. Biol. Cell 8 (9), 1763-1775 (1997)
   PUBMED   9307972
REFERENCE   2  (bases 1 to 3576)
  AUTHORS   Escalante,R., Wessels,D., Soll,D. and Loomis,W.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1997) Department of Biology, University of
            California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-032,
            USA
FEATURES             Location/Qualifiers
     source          1. .3576
                     /organism="Dictyostelium discoideum"
                     /mol_type="genomic DNA"
                     /strain="AX4"
                     /sub_strain="AK244"
                     /db_xref="taxon:44689"
                     /chromosome="4"
                     /map="4.25"
     gene            884. .3423
                     /gene="migA"
     mRNA            join(<884. .2899,2998. .>3423)
                     /gene="migA"
                     /product="MigA"
     CDS             join(884. .2899,2998. .3423)
                     /gene="migA"
                     /function="essential for slug migration"
                     /note="N-terminus of this protein is similar to other BTB
                     domain proteins, such as the Mus musculus and human zinc
                     finger proteins encoded by GenBank Accession Numbers
                     U14556 and L16896, respectively"
                     /codon_start=1
                     /product="MigA"
                     /protein_id="AAB47544.1"
                     /db_xref="GI:1841872"
                     /translation="MATINIINNNNNQNNNNQNNNNNINHTEYFNNGFKSLLGEFKKY
                     TDIKIKCGGQIFPTHRIILANSSEFFGRLLLSDFKESSQSVIELKQPDPFNVFPMVLE
                     FMYDGIIYLSPENVIPLLAMSDHYLIRDLQHLCLLFLDQNLHRENVLTILKYSIDFHF
                     EPITERCITIISKNFTFVCENLYKQTKINQNNNNNCINIENNKKHLNTSSASSSLNSS
                     TSSASSSSSSNNNYYTTTTTTTTTTASSSNFNKNLTNCKVTGQQQQQQQQQQQQQQ
                     QQQQQQQQQQPIYDFSFLPASIFYQILDHDYLVVNSEYTLYRIVLNFIEKFFGENFTE
                     ITYNTNITDQLKVSDLMSQIRYILMPYDQLTLITKDPLIPKEILFETLLERLKRFEGK
                     TSSPSSSSTSSSPSNTSNSPTKQTSPSKYQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
                     PQQQPQQQPQQPQPQIIQQQPQQPQIIQQQPQQLQQLQQPQQPQTIQQMVQQQQNIQQ
                     QQQQQQQQQNIQQQQTQNMQQQLQQQLLQQIQQQQQQQQQPQIIKQFTPSSSPLSSSF
                     ISLNASQSKLSINNATNTTNNNNNNSNNVTNTNNNNIINNNNNNNNKQTQSNNNDNKI
                     LKRYQPRPPQSILFEYSFDFDFKGIIFWISTDGGNEKWSNPHSTSKIKITSSSIDKGN
                     LYDIVELTPNAFWTKDVPASWVMIDLGPNRTVVPMYYTIRHGLSYKSDSLRTWDFQGS
                     TNGEQWTVLKRHTNDPSLNYKYATHSWPVTGCETAFRYFRILQTGKNSNNRNFLVIGG
                     LEIYGELCETNPNPN"
     exon            <884. .2899
                     /gene="migA"
                     /number=1
     misc_feature    951. .1289
                     /gene="migA"
                     /note="encodes BTB domain; similar to the Drosophila
```

```
                melanogaster BTB domain encoded by GenBank Accession
                Number U01333"
     variation  1253^1254
                /gene="migA"
                /note="gene interruption in mutant"
                /replace="pbsr1 external plasmid, approximately 4 kb".
     intron     2900. .2997
                /gene="migA"
                /number=1
     exon       2998. .>3423
                /gene="migA"
                /number=2
ORIGIN

  Query Match             5.4%;  Score 71.2;  DB 2;  Length 3576;
  Best Local Similarity   47.2%;  Pred. No. 0.0049;
  Matches  217;  Conservative    0;  Mismatches  243;  Indels    0;  Gaps    0;

Qy        412 TGTCTAGTGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAAC 471
              | | || | |    ||||   || |     |      |  | ||||||  ||  | |  ||
Db        852 TATTTATTATTATTTATTATTACAATTTTGTTTTTGTATAAAAAAATAAAAAATAAAAAA 793

Qy        472 TAATCCACTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCT 531
              |||   |    |||   |  ||  ||  |  |||  || | | |  |||   |   |
Db        792 TAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAACAAACAAATTAAATTTAA 733

Qy        532 CTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAG 591
              ||||||||  |  | ||    ||  | | ||| ||   || |   ||  ||     ||
Db        732 TTTAAAAAAATAAAAAATAAAAAATAAAATAAAATAAAATAAAAAATAATAATTAAAAAA 673

Qy        592 AGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGT 651
              |   | |||  | |   |||  || | | ||       |   |||| |||||  ||  |
Db        672 AAAAAAAAAAAAAAAATAAATATAGGGTGTGTGAAGAATTGTTTTTTTTTTTGTTTGTTG 613

Qy        652 TATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGA 711
              |  ||  || |  | ||| || | | ||| ||      |  | | |    ||      |
Db        612 TTGTTATTTTTGTTATTTTTAAATAATTTATTTATACAATAAAAATACTTTAAACACAAC 553

Qy        712 TTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGT 771
                |  | |  || ||    |  ||||| | |       |  | |  |  ||   | |
Db        552 ACACAACACACCACAATTGTTTTTTTTTTGTGTTTTTTTGTGTTTTTTTTTTATCAACCT 493

Qy        772 TCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAAT 831
              |  ||    | || ||||||  ||       |  |||     |     ||| |||  || |||
Db        492 TTCAATCGCCTTTTGATTTAGTTTTTTTTTTTTTTTTTGAATTGTTTTTTTTTGAAAAAT 433

Qy        832 TTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTAT 871
              ||  |||| |  ||    |||  |||||||| | ||| |
Db        432 ATTTTTTTTATTTAATTTTATTTTTTTATTTTATTTTTTT 393

Search completed: March  5, 2006, 21:55:27
Job time : 5725.47 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 18:11:59 ; Search time 6189.32 Seconds
(without alignments)
9910.279 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311
Perfect score: 1311
Sequence: 1 ccatggtgtctctatgaaaa.....cgcacatatgtttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	908	69.3	963	9	CC656933	CC656933 OGWDQ20TM
	2	679	51.8	915	10	CG224225	CG224225 OG1AG08TV
c	3	419.8	32.0	687	9	CC656939	CC656939 OGWDQ20TV
c	4	96	7.3	715	10	CG252571	CG252571 OG4BB05TC
	5	92	7.0	967	10	CL235046	CL235046 ZMMBBb057
	6	91	6.9	860	10	CZ295176	CZ295176 ZMMBF0063
	7	88	6.7	814	10	CG048704	CG048704 PUILO19TB
c	8	87.8	6.7	754	10	CG414922	CG414922 ZMMBBb029
c	9	86.8	6.6	950	9	CC439901	CC439901 PUHRV15TB
	10	86.6	6.6	652	9	CC384247	CC384247 PUHOC67TB
c	11	86.6	6.6	797	9	CC400575	CC400575 PUHLU61TD
	12	86.6	6.6	820	9	CC400574	CC400574 PUHLU61TB
	13	85.4	6.5	471	10	CG103452	CG103452 PUJBE19TB
c	14	85.2	6.5	765	10	CG082135	CG082135 PUFQX12TD
c	15	85.2	6.5	781	9	CC630219	CC630219 OGUCG53TV
c	16	85.2	6.5	815	10	CG349565	CG349565 OG0FI83TH
	17	85.2	6.5	834	9	CC630210	CC630210 OGUCG53TH
	18	85	6.5	789	9	CC433618	CC433618 PUHHP17TD
c	19	84.6	6.5	1092	10	CNS020K7	AL175696 Tetraodon
	20	84	6.4	793	9	BZ816381	BZ816381 PUFBA86TD
c	21	84	6.4	1078	10	CL997678	CL997678 ZMMBHf001
	22	83.8	6.4	530	10	CG201774	CG201774 PUICH24TB
c	23	83.6	6.4	722	10	CG333914	CG333914 OG0AD14TH
	24	83.6	6.4	722	10	CG333929	CG333929 OG0AD14TV
	25	83.6	6.4	781	10	CG034985	CG034985 PUIGR68TB
	26	83.6	6.4	861	9	BZ797976	BZ797976 PUGBP28TB
c	27	83.6	6.4	865	9	CC430754	CC430754 PUHEP05TB
c	28	81.8	6.2	861	10	CG102092	CG102092 PUFYW94TD
c	29	81.8	6.2	947	9	CC435780	CC435780 PUHSN02TD
	30	81.8	6.2	981	9	BZ784278	BZ784278 PUFHV09TD
c	31	81.8	6.2	1017	9	CC620594	CC620594 OGUCC26TV
	32	81.6	6.2	649	9	CC613918	CC613918 OGJAG88TV
c	33	81.6	6.2	733	9	BZ778636	BZ778636 ih02f10.g
c	34	81.6	6.2	925	10	CG071791	CG071791 PUJBK02TB
c	35	81.6	6.2	960	9	BZ676889	BZ676889 PUBIG17TD
	36	81.6	6.2	994	9	CC003943	CC003943 PUDJN65TD
	37	81.6	6.2	1016	10	CL996481	CL996481 ZMMBHf000
	38	81.6	6.2	1022	10	CL984151	CL984151 ZMMBHd000
	39	81.4	6.2	999	9	CC385762	CC385762 PUHFK14TD
	40	81.2	6.2	756	10	CG221693	CG221693 OGWMG04TH
c	41	81	6.2	1101	10	CNS00LT2	AL078714 Drosophil
c	42	80.8	6.2	792	8	DR794756	DR794756 ZM_BFb001
	43	80.6	6.1	699	9	BZ996930	BZ996930 PUGIR50TB
c	44	80.6	6.1	832	9	CC676387	CC676387 OGKAS81TH
c	45	80.6	6.1	839	10	CG246849	CG246849 OGWAL22TH

ALIGNMENTS

RESULT 1
CC656933
LOCUS CC656933 963 bp DNA linear GSS 19-JUN-2003
DEFINITION OGWDQ20TM ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15, genomic survey sequence.
ACCESSION CC656933
VERSION CC656933.1 GI:32060225
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 963)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGWDQ20TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: methylation filtered.
FEATURES Location/Qualifiers
source 1. .963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0554D15"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 69.3%; Score 908; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 9.6e-180;
Matches 908; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 404 ACCAGCCATGTCTAGTGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACC 463

```
Db          1 ACCAGCCATGTCTAGTGTCCACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACC 60

Qy        464 AAAGTAACTAATCCACTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACC 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGTAACTAATCCACTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACC 120

Qy        524 ACGATCCTCTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAA 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACGATCCTCTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAA 180

Qy        584 CATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTC 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTC 240

Qy        644 ATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTT 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTT 300

Qy        704 TCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAG 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAG 360

Qy        764 ATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTT 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTT 420

Qy        824 CAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAAC 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAAC 480

Qy        884 ACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCA 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCA 540

Qy        944 CTGTGTATAATTATTTGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGA 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGTGTATAATTATTTGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGA 600

Qy       1004 AACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAG 660

Qy       1064 TTTTAAGTCATTGTCCCTGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCT 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTTAAGTCATTGTCCCTGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCT 720

Qy       1124 CATAGGTGCCTACGTCAGGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGC 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CATAGGTGCCTACGTCAGGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGC 780

Qy       1184 TTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGAT 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGAT 840

Qy       1244 ACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTT 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACCTACTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTT 900

Qy       1304 TCCTGAAC 1311
              ||||||||
Db        901 TCCTGAAC 908
```

RESULT 2
CG224225

```
LOCUS       CG224225                915 bp    DNA     linear   GSS 22-AUG-2003
DEFINITION  OG1AG08TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0716B15,
            genomic survey sequence.
ACCESSION   CG224225
VERSION     CG224225.1  GI:34124113
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG1AG08TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0716B15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN
```

Query Match 51.8%; Score 679; DB 10; Length 915;
Best Local Similarity 99.9%; Pred. No. 9.5e-132;
Matches 690; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy        621 TGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTT 680
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db          1 TGACGTTCAGATTTTTCTTTTTC-TTCTTGTTATTTTGTTATTGTTTTTATATACATTTT 59

Qy        681 CTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTT 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 CTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTT 119

Qy        741 ATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGC 800
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 ATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGC 179

Qy        801 TAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTAT 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTAT 239

Qy        861 TTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAAT 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAAT 299

Qy        921 GTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTGGCCAGCCCCATAAATTATTT 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTGGCCAGCCCCATAAATTATTT 359

Qy        981 AAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAG 1040
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 AAACCGAAACTGAAATCGAGCGAAACCAAATCTGAGCTATTTCTCTAGATTAGTAAAAAG 419

Qy       1041 GGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCCTGAGATGTGCGGTTTGGCAA 1100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 GGAGAGAGAGAGGAAGAAATCAGTTTTAAGTCATTGTCCCTGAGATGTGCGGTTTGGCAA 479

Qy       1101 CGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCAGGTTCGGCAGCTCTCGTGTC 1160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CGATAGCCACCGTAATCATAGCTCATAGGTGCCTACGTCAGGTTCGGCAGCTCTCGTGTC 539

Qy       1161 ATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCT 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        540 ATCTCACATGGCATACTACATGCTTGTTCAACCGTTCGTCTTGTTCCATCGTCCAAGCCT 599

Qy       1221 TGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTT 1280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TGCCTATTCTGAACCAAGAGGATACCTACTCCCAAACAATCCATCTTACTCATGCAACTT 659

Qy       1281 CCATGCAAACACGCACATATGTTTCCTGAAC 1311
             |||||||||||||||||||||||||||||||
Db        660 CCATGCAAACACGCACATATGTTTCCTGAAC 690


RESULT 3
CC656939/c
LOCUS       CC656939                687 bp    DNA     linear   GSS 19-JUN-2003
DEFINITION  OGWDQ20TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0554D15,
            genomic survey sequence.
ACCESSION   CC656939
VERSION     CC656939.1  GI:32060231
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 687)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OGWDQ20TM
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .687
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0554D15"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match           32.0%;  Score 419.8;  DB 9;  Length 687;
  Best Local Similarity   99.5%;  Pred. No. 2e-77;
  Matches  421;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        889 GAAAAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTG 948
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GAAAAATAGAGTGCCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTG 628

Qy        949 TATAATTATTTGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCA 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TATAATTATTTGGCCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCA 568

Qy       1009 AATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTA 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 AATCTGAGCTATTTCTCTAGATTAGTAAAAAGGGAGAGAGAGAGGAAGAAATCAGTTTTA 508

Qy       1069 AGTCATTGTCCCTGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAG 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 AGTCATTGTCCCTGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAG 448

Qy       1129 GTGCCTACGTCAGGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTT 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        447 GTGCCTACGTCAGGTTCGGCAGCTCTCGTGTCATCTCACATGGCATACTACATGCTTGTT 388

Qy       1189 CAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTA 1248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CAACCGTTCGTCTTGTTCCATCGTCCAAGCCTTGCCTATTCTGAACCAAGAGGATACCTA 328

Qy       1249 CTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTG 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CTCCCAAACAATCCATCTTACTCATGCAACTTCCATGCAAACACGCACATATGTTTCCTG 268

Qy       1309 AAC 1311
              |||
Db        267 AAC 265


RESULT 4
CG252571/c
LOCUS       CG252571                715 bp    DNA     linear   GSS 25-AUG-2003
DEFINITION  OG4BB05TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0809B10,
            genomic survey sequence.
ACCESSION   CG252571
VERSION     CG252571.1  GI:34154661
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 715)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .715
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0809B10"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match           7.3%;  Score 96;  DB 10;  Length 715;
  Best Local Similarity   80.1%;  Pred. No. 1.5e-09;
  Matches  125;  Conservative    0;  Mismatches   30;  Indels    1;  Gaps    1;

Qy        135 TCTCCCACCTCGCGTTGCACATTTCTTGGATGTCGGTGGTTCCCATCTGACCGAGGCCCA 194
              || | ||||     ||| | ||||||||||||||| ||||||| ||||| |||||||||
Db        356 TCACACACCCGCTCTTGTGCCTTTCTTGGATGTCGGCGGTTCCCGTCTGATCGAGGCCCA 297

Qy        195 TCAGACACCTTTCGGGACACCCATCAAGGGCCTTTCGGATGGCCCACGAGACGTATCGGG 254
              | || ||||| ||||   |||||| || ||||||||||||||||||| ||| |   |||
Db        296 TAAGGCACCTGTCGGAGAACCCATTAAAGGCCTTTCGGATGGCCCACGGGACAT-CGGGG 238

Qy        255 TCGTGGTGATCCAGGGGATATATGTCCCCCACAATC 290
              ||||||||| || |||||||| | |||||||||| |
Db        237 TCGTGGTGACCCGGGGGATATCTATCCCCCACAACC 202


RESULT 5
CL235046
```

```
LOCUS       CL235046                967 bp    DNA     linear   GSS 15-JAN-2004
DEFINITION  ZMMBBb0575O01r ZMMBBb (HindIII) Zea mays genomic clone
            ZMMBBb0575O01 3', genomic survey sequence.
ACCESSION   CL235046
VERSION     CL235046.1  GI:40891729
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 967)
  AUTHORS   Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
            Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
  TITLE     Sequencing of the maize genome at PGIR (2003c)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Bharti,A.K.
            Dr.Joachim Messing's lab
            The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
            University
            190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            Tel: 732 445 3801
            Fax: 732 445 5735
            Email: bharti@waksman.rutgers.edu
            Seq primer: SP6
            Class: BAC ends
            High quality sequence start: 64.
FEATURES             Location/Qualifiers
     source          1. .967
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0575O01"
                     /lab_host="E. coli DH10B"
                     /clone_lib="ZMMBBb (HindIII)"
                     /note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"
ORIGIN

  Query Match             7.0%;  Score 92;  DB 10;  Length 967;
  Best Local Similarity   78.6%;  Pred. No. 1e-08;
  Matches  110;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy        150 TGCACATTTCTTGGATGTCGGTGGTTCCCATCTGACCGAGGCCCATCAGACACCTTTCGG 209
              ||| | ||||||||||  ||| ||||||| |||||  ||||||||||   ||||||| ||
Db        827 TGCGCCTTTCTTGGATACCGGCGGTTCCCGTCTGATTGAGGCCCATCCAGCACCTTTTGG 886

Qy        210 GACACCCATCAAGGGCCTTTCGGATGGCCCACGAGACGTATCGGGTCGTGGTGATCCAGG 269
              |   ||||||||| |||||||||||||| ||||| || ||||| || |   ||||  |  ||
Db        887 GTGGCCCATCAAGTGCCTTTCGGATGGTCCACGGGATGTATCCGGGCCGTGTGACCTGGG 946

Qy        270 GGATATATGTCCCCCACAAT 289
              ||||||| |||||||||| | |
Db        947 GGATATCTGTCCCCCATAGT 966


RESULT 6
CZ295176
LOCUS       CZ295176                860 bp    DNA     linear   GSS 23-MAR-2005
DEFINITION  ZMMBF0063M20f ZMMBF Zea mays genomic clone ZMMBF0063M20 5', genomic
            survey sequence.
ACCESSION   CZ295176
VERSION     CZ295176.1  GI:61708952
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 860)
  AUTHORS   Bharti,A.K., Nelson,A.B., Young,S., Keizer,G., Zohovetz,V., Fuks,G.
            and Messing,J.
  TITLE     Construction, Sequencing and Characterization of a Fosmid Library
            of the B73 Maize Genome
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Bharti,A.K.
            Dr.Joachim Messing's lab
            The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
            University
            190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            Tel: 732 445 3801
            Fax: 732 445 5735
            Email: bharti@waksman.rutgers.edu
            Seq primer: Fos_F
            Class: fosmid ends.
FEATURES             Location/Qualifiers
     source          1. .860
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBF0063M20"
                     /lab_host="EPI100-T1"
                     /clone_lib="ZMMBF"
                     /note="Vector: pEpiFOS-5; Site_1: Eco72I"
ORIGIN

  Query Match             6.9%;  Score 91;  DB 10;  Length 860;
  Best Local Similarity   62.6%;  Pred. No. 1.7e-08;
  Matches  142;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy        782 TTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTT 841
              |||  |    | |||| ||||||||| ||||||| ||| ||  || ||||||| ||||||
Db         51 TTCAAAGCACTCTAAGGGCTAGTTTGACAACCCTATTTTTTCTAAGGAATTTTTATTTTT 110

Qy        842 TCAAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTT 901
              | || |||||||||||| |||||||||    ||||| |||| | |||| ||||||| ||||
Db        111 TAAAGAAAAATTAGTTCATTTTCTCTTGTGAAAATGGAAATCCCTTAAAAAAATAAAGTT 170

Qy        902 GCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTGG 961
              | | | ||||| || ||  |  |   |  |    | ||| |  ||| || |  |   |
Db        171 ACTAAATTAGCCTTAAAAGTTGGTTTGATGACCACAAAATTAGAGTGGATTAAGAGACGA 230

Qy        962 CCAGCCCCATAAATTATTTAAACCGAAACTGAAATCGAGCGAAACCA 1008
                |  |||  |       |  |  |  |  ||| |||     | |||
Db        231 GAAATCCCTTGCTAATGATTGAGGGGGAGAGAATTCGTAACTATCCA 277


RESULT 7
CG048704
LOCUS       CG048704                814 bp    DNA     linear   GSS 19-AUG-2003
DEFINITION  PUILO19TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0611C13,
            genomic survey sequence.
ACCESSION   CG048704
VERSION     CG048704.1  GI:33920884
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 814)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUILO19TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
```

```
FEATURES             Location/Qualifiers
     source          1. .814
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa0611C13"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match           6.7%;  Score 88;  DB 10;  Length 814;
  Best Local Similarity  62.7%;  Pred. No. 7.1e-08;
  Matches  153;  Conservative    0;  Mismatches   90;  Indels    1;  Gaps    1;

Qy        713 TTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTT 772
              |||| | | |    || ||         |||||     |  || |  ||||   | |
Db        487 TTTAGAGGACGTTGGTAGAGATGGAAAGATATAGAGAAAATAATCTTTTAGAGAATGTTA 546

Qy        773 CAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATT 832
              |||   |          ||| ||||||||||||||||| || ||  ||| |||||||| |||
Db        547 TAAAGGACAAGAAACATTTCTTAAGAGCTAGTTTGGTAATCCCATTTTTTTCAAAGGATT 606

Qy        833 TTGATTTTTTCAAAA-AAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAA 891
              || ||| |  |||   |||| ||||||||| |  ||| |||||| | || ||||  |||
Db        607 TTCATTCTCACAAGGGAAAAATAGTTTATTTCCCTTTTGTAAAATGGTAATCACTCGGAA 666

Qy        892 AAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTAT 951
              ||||||||||  || ||||||||||   |  | |||| |||     | ||| |  | | |
Db        667 AAATAGAGTTCTCAAACTAGCCCTAAGGTTATATCCCGATAATCCCCTATCTCATTCTCT 726

Qy        952 AATT 955
              |||
Db        727 ATTT 730


RESULT 8
CG414922/c
LOCUS       CG414922                754 bp    DNA     linear   GSS 08-SEP-2003
DEFINITION  ZMMBBb0290L09.r ZMMBBb Zea mays genomic clone ZMMBBb0290L09 3',
            genomic survey sequence.
ACCESSION   CG414922
VERSION     CG414922.1  GI:34505144
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 754)
  AUTHORS   Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
            and Wing,R.
  TITLE     Sequencing of the maize genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: rwing@genome.arizona.edu
            PCR PRimers
            FORWARD: T7
            BACKWARD: M13r
            Plate: 0290  row: L  column: 09
            Seq primer: M13r
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .754
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0290L09"
                     /lab_host="DH10B"
                     /clone_lib="ZMMBBb"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII; Zea mays L. ssp. mays"
ORIGIN

  Query Match           6.7%;  Score 87.8;  DB 10;  Length 754;
  Best Local Similarity  71.2%;  Pred. No. 7.8e-08;
  Matches  116;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy        779 TCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATT 838
              |||   |  | ||  |||  |||||||||  || |   ||| |||| ||| ||||| |||
Db        680 TCTAAATTTTATTCCTAAAGGCTAGTTTGAAAAGCACATTTTTTTTCCAAGGATTTTAATT 621

Qy        839 TTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGA 898
              |||  ||| |||||||| || ||||| |||| | ||||||||||  ||| |||||||||| |
Db        620 TTTCTAAATAAAATTAATTCATTTTTTCTTGAAAAAATAGAAATCACATAGAAAAATAAA 561

Qy        899 GTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAAT 941
              ||| | | |||| | |||||  | |||    | | ||||| |
Db        560 GTTCCTAAACTAACTCTAAAAAATATTCTTCGTCATTTACAGT 518


RESULT 9
CC439901/c
LOCUS       CC439901                950 bp    DNA     linear   GSS 20-MAY-2003
DEFINITION  PUHRV15TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa519D06,
            genomic survey sequence.
ACCESSION   CC439901
VERSION     CC439901.1  GI:30940477
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 950)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUHRV15TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .950
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa519D06"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match           6.6%;  Score 86.8;  DB 9;  Length 950;
  Best Local Similarity  72.4%;  Pred. No. 1.3e-07;
  Matches  126;  Conservative    0;  Mismatches   47;  Indels    1;  Gaps    1;

Qy        787 ATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAA 846
              || ||| |||   |||||||| |||| | ||| | |||  | ||||| |||||||||
Db        241 ATCTTTCTAATGACTAGTTTGGGAACCTTATTT-TCTCACGGGATTTTTATTTTTCAAT 183
```

```
Qy        847 AAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAG 906
              |||||||||||||||||||| | |||||| ||| |||||   |||||| |||| |||
Db        182 GAAAATTAGTTTATTTTCTCTTGAGAAAATATAAATCACTTGAGAAAATATAGTTTCCAA 123

Qy        907 ACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              |||||||||| | | ||||   ||  |   | ||     || |||  ||  |||
Db        122 ACTAGCCCTAAAGTATTTTATAAAATAGAGAAAAAATTCGTCTATTTTTCATTG 69


RESULT 10
CC384247
LOCUS       CC384247                652 bp    DNA     linear   GSS 19-MAY-2003
DEFINITION  PUHOC67TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa495L13,
            genomic survey sequence.
ACCESSION   CC384247
VERSION     CC384247.1  GI:30864143
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 652)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUHOC67TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa495L13"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             6.6%;  Score 86.6;  DB 9;  Length 652;
  Best Local Similarity   63.9%;  Pred. No. 1.4e-07;
  Matches  131;  Conservative    0;  Mismatches   74;  Indels    0;  Gaps    0;

Qy        714 TTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTC 773
              | || || ||      ||  |  ||  ||   ||      |||| ||  |||   | |
Db        341 TAATTAATTGTAGTACTAGCTACATTCTTAATCGACATATTATGTAGTTAATTAATGCTA 400

Qy        774 AAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTT 833
              | ||   | | |  |  | |||||||||||||||| |||  | ||| | | | ||| |||
Db        401 AGAATCTTATTTATTAATGTTAAGAGCTAGTTTGGGAACTATATTTTTCTAAGAGATTTT 460

Qy        834 TGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAA 893
              |  |||||||||   |||||||||||||||  | |   ||||||||||  | || |||||
Db        461 TATTTTTTTCAAGTGAAATTAGTTTATTTTTCCATGGGAAAATAGAAATTAATTGGAAAA 520

Qy        894 ATAGAGTTGCCAGACTAGCCCTAGA 918
              ||   ||| ||| ||||||||||
Db        521 ATTAGGTTTCCAAACTAGCCCTAAA 545


RESULT 11
CC400575/c
```

```
LOCUS       CC400575                797 bp    DNA     linear   GSS 19-MAY-2003
DEFINITION  PUHLU61TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa480L01,
            genomic survey sequence.
ACCESSION   CC400575
VERSION     CC400575.1  GI:30880665
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 797)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUHLU61TB
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .797
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa480L01"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             6.6%;  Score 86.6;  DB 9;  Length 797;
  Best Local Similarity   59.7%;  Pred. No. 1.4e-07;
  Matches  163;  Conservative    0;  Mismatches  109;  Indels    1;  Gaps    1;

Qy        713 TTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTT 772
              |||| |  | |   || ||      |||||    |   || |  ||||   | |
Db        577 TTTAGAGGACGTTGGTAGAGATGGAAAGATATAGAGAAAATAATCTTTTAGAGAATGTTA 518

Qy        773 CAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATT 832
              |||   |        |||||||||||||||||||| || ||  ||| |||||||| |||
Db        517 TAAAGGACAAGAAACATTTTTTAAGAGCTAGTTTGGTAATCCCATTTTTTTCAAAGGATT 458

Qy        833 TTGATTTTTTCAAAA-AAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAA 891
              || ||| |  |||   |||| |||||||||| |  ||| |||||| | || ||||  |||
Db        457 TTCATTCTCACAAGGGAAAAATAGTTTATTTCCCTTTTGTAAAATGGTAATCACTCGGAA 398

Qy        892 AAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTAT 951
              |||||||||  || ||||||||||   | | ||| |  ||    | ||| |  | | |
Db        397 AAATAGAGTTCTCAAACTAGCCCTAAGGTTATATCCGATAAATCCCCTATCTCATTCTCT 338

Qy        952 AATTATTTGGCCAGCCCCATAAATTATTTAAAC 984
              | ||      |||       ||||  | | | |
Db        337 ATTTCAAATTCCACTTTGCAAAATAGTGTCATC 305


RESULT 12
CC400574
LOCUS       CC400574                820 bp    DNA     linear   GSS 19-MAY-2003
DEFINITION  PUHLU61TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa480L01,
            genomic survey sequence.
ACCESSION   CC400574
VERSION     CC400574.1  GI:30880664
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
```

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 820)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUHLU61TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .820
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa480L01"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             6.6%;  Score 86.6;  DB 9;  Length 820;
  Best Local Similarity   59.7%;  Pred. No. 1.4e-07;
  Matches  163;  Conservative    0;  Mismatches  109;  Indels    1;  Gaps    1;

Qy        713 TTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTT 772
              |||| |  | |     || ||         |||||      |   || |  ||||   | |
Db        374 TTTAGAGGACGTTGGTAGAGATGGAAAGATATAGAGAAAATAATCTTTTAGAGAATGTTA 433

Qy        773 CAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATT 832
              |||    |         |||||||||||||||||||||| || ||  ||| |||||||| |||
Db        434 TAAAGGACAAGAAACATTTTTTAAGAGCTAGTTTGGTAATCCCATTTTTTTCAAAGGATT 493

Qy        833 TTGATTTTTTCAAAA-AAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAA 891
              || ||| |  |||   |||| |||||||| |  ||| ||||||  | || ||||  |||
Db        494 TTCATTCTCACAAGGGAAAAATAGTTTATTTCCCTTTTGTAAAATGGTAATCACTCGGAA 553

Qy        892 AAATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTAT 951
              ||||||||||  || ||||||||||   |  | ||| |  ||    | ||| |  | | |
Db        554 AAATAGAGTTCTCAAACTAGCCCTAAGGTTATATCCGATAAATCCCCTATCTCATTCTCT 613

Qy        952 AATTATTTGGCCAGCCCCATAAATTATTTAAAC 984
              | ||     |||      ||||  | | | |
Db        614 ATTTCAAATTCCACTTTGCAAAATAGTGTCATC 646


RESULT 13
CG103452
LOCUS       CG103452                471 bp    DNA     linear   GSS 20-AUG-2003
DEFINITION  PUJBE19TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0628D13,
            genomic survey sequence.
ACCESSION   CG103452
VERSION     CG103452.1  GI:33985746
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 471)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUJBE19TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .471
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBTa0628D13"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                     CoT selected genomic DNA library"
ORIGIN

  Query Match             6.5%;  Score 85.4;  DB 10;  Length 471;
  Best Local Similarity   69.5%;  Pred. No. 2.4e-07;
  Matches  116;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy        789 TTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAA 848
              |||| | |||||||||||||| || ||  ||| ||  ||  || ||     |||||||||
Db        216 TTTTCTTAGAGCTAGTTTGGAAATCCCATTTTTTCCAGGGATTTACATTTTTTTCAAGGG 275

Qy        849 AAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGAC 908
              ||||||||| ||||||  || | |||||||| || ||| | |  |||||| ||| ||| ||
Db        276 AAATTAGTTCATTTTCCTTTGAAAAAATAGGAATCACATGGGGAAATAGTGTTCCCAAAC 335

Qy        909 TAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATT 955
              |||| ||    |   ||| | || ||| || |  || ||. ||  |||
Db        336 TAGCGCTTAGACTATTTTCTAAAAAAGTAGATCCAATGAATAAGATT 382


RESULT 14
CG082135/c
LOCUS       CG082135                765 bp    DNA     linear   GSS 20-AUG-2003
DEFINITION  PUFQX12TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0712B24,
            genomic survey sequence.
ACCESSION   CG082135
VERSION     CG082135.1  GI:33964429
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
  TITLE     Maize Genomics Consortium
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: PUFQX12TB
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
```

```
                    /clone="ZMMBTa0712B24"
                    /clone_lib="ZM_0.6_1.0_KB"
                    /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                    CoT selected genomic DNA library"
ORIGIN

  Query Match             6.5%;  Score 85.2;  DB 10;  Length 765;
  Best Local Similarity   71.8%;  Pred. No. 2.7e-07;
  Matches  125;  Conservative    0;  Mismatches   48;  Indels    1;  Gaps    1;

Qy        787 ATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAA 846
              || ||| |||   |||||||||| |||| | ||| | |||  | ||||| ||||||||||
Db        255 ATCTTTCTAATGACTAGTTTGGGAACCTTATTT-TCTCACGGGATTTTTATTTTTTCAAT 197

Qy        847 AAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAG 906
              |||||||||||||||||||| | |||||| ||| |||||   |||||| |||| |||
Db        196 GAAAATTAGTTTATTTTCTCTTGAGAAAATATAAATCACTTGAGAAAATATAGTTTCCAA 137

Qy        907 ACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              |||||||||| | | ||||   ||  |     ||     || |||  ||  |||
Db        136 ACTAGCCCTAAAGTATTTTATAAAATAGAGGAAAAATTCGTCTATTTTTCATTG 83


RESULT 15
CC630219/c
LOCUS       CC630219                781 bp    DNA     linear   GSS 19-JUN-2003
DEFINITION  OGUCG53TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0406J09,
            genomic survey sequence.
ACCESSION   CC630219
VERSION     CC630219.1  GI:32003354
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 781)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OGUCG53TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1. .781
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0406J09"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             6.5%;  Score 85.2;  DB 9;  Length 781;
  Best Local Similarity   71.8%;  Pred. No. 2.7e-07;
  Matches  125;  Conservative    0;  Mismatches   48;  Indels    1;  Gaps    1;

Qy        787 ATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAA 846
              || ||| |||   |||||||||| |||| | ||| | |||  | ||||| ||||||||||
Db        665 ATCTTTCTAATGACTAGTTTGGGAACCTTATTT-TCTCACGGGATTTTTATTTTTTCAAT 607

Qy        847 AAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAG 906
              |||||||||||||||||||| | |||||| ||| |||||   |||||| |||| |||

Db        606 GAAAATTAGTTTATTTTCTCTTGAGAAAATATAAATCACTTGAGAAAATATAGTTTCCAA 547

Qy        907 ACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTG 960
              |||||||||| | | ||||   ||  |     ||     || |||  ||  |||
Db        546 ACTAGCCCTAAAGTATTTTATAAAATAGAGGAAAAATTCGTCTATTTTTCATTG 493


Search completed: March  6, 2006, 01:57:38
Job time : 6193.32 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 01:58:03 ; Search time 1368.54 Seconds
(without alignments)
2100.307 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311
Perfect score: 1311
Sequence: 1 ccatggtgtctctatgaaaa.....cgcacatatgtttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	68	5.2	1092	11	US-11-014-071-2	Sequence 2, Appli
2	57.8	4.4	9347	8	US-10-240-708-36	Sequence 36, Appl
3	55.6	4.2	254396	7	US-10-330-773-534	Sequence 534, App
4	54.8	4.2	915	6	US-09-925-065A-54272	Sequence 54272, A
5	53.6	4.1	488	6	US-09-925-065A-803046	Sequence 803046,
c 6	53.6	4.1	5152	8	US-10-240-708-73	Sequence 73, Appl
7	53.4	4.1	5562	8	US-10-240-708-63	Sequence 63, Appl
c 8	53.2	4.1	119036	8	US-10-995-561-13314	Sequence 13314, A
9	52	4.0	488	6	US-09-925-065A-803045	Sequence 803045,
c 10	51.6	3.9	6317	8	US-10-240-708-11	Sequence 11, Appl
c 11	51.6	3.9	355211	7	US-10-330-773-242	Sequence 242, App
12	51.4	3.9	173602	12	US-11-121-086-25	Sequence 25, Appl
13	51.2	3.9	687411	7	US-10-330-773-26	Sequence 26, Appl
14	51	3.9	677	6	US-09-925-065A-667262	Sequence 667262,
15	51	3.9	677	6	US-09-925-065A-667263	Sequence 667263,
16	51	3.9	677	6	US-09-925-065A-667265	Sequence 667265,
c 17	50.6	3.9	533	6	US-09-925-065A-13297	Sequence 13297, A
c 18	50.6	3.9	610	6	US-09-925-065A-297109	Sequence 297109,
c 19	50.6	3.9	4330	12	US-11-091-883-182	Sequence 182, App
c 20	50.2	3.8	583	6	US-09-925-065A-206148	Sequence 206148,
21	49.8	3.8	677	6	US-09-925-065A-667264	Sequence 667264,
22	49.8	3.8	677	6	US-09-925-065A-667266	Sequence 667266,
c 23	49.8	3.8	4339	8	US-10-909-125-801	Sequence 801, App
c 24	49	3.7	583	6	US-09-925-065A-206150	Sequence 206150,
25	49	3.7	10467	8	US-10-240-708-2	Sequence 2, Appli
c 26	48.6	3.7	583	6	US-09-925-065A-206149	Sequence 206149,
27	48.2	3.7	5152	8	US-10-240-708-74	Sequence 74, Appl
28	48.2	3.7	8467	12	US-11-011-332A-96	Sequence 96, Appl
29	48.2	3.7	355211	7	US-10-330-773-242	Sequence 242, App
30	48	3.7	6499	12	US-11-011-332A-76	Sequence 76, Appl
31	48	3.7	108057	7	US-10-330-773-835	Sequence 835, App
c 32	47.6	3.6	6113	8	US-10-240-708-13	Sequence 13, Appl
c 33	47.4	3.6	173602	12	US-11-121-086-25	Sequence 25, Appl
34	47.2	3.6	6070	8	US-10-240-708-10	Sequence 10, Appl
35	47.2	3.6	10619	8	US-10-240-708-3	Sequence 3, Appli
36	47	3.6	19233	8	US-10-240-708-46	Sequence 46, Appl
37	46.8	3.6	8537	8	US-10-240-708-42	Sequence 42, Appl
38	46.8	3.6	8961	8	US-10-240-708-80	Sequence 80, Appl
39	46.6	3.6	9347	8	US-10-240-708-35	Sequence 35, Appl
c 40	46.6	3.6	86899	7	US-10-330-773-963	Sequence 963, App
41	46.4	3.5	6499	12	US-11-011-332A-90	Sequence 90, Appl
c 42	46.2	3.5	3219	6	US-09-925-065A-714059	Sequence 714059,
c 43	46.2	3.5	3219	6	US-09-925-065A-714060	Sequence 714060,
44	46.2	3.5	6499	12	US-11-011-332A-77	Sequence 77, Appl
45	46.2	3.5	6499	12	US-11-011-332A-91	Sequence 91, Appl

ALIGNMENTS

RESULT 1
US-11-014-071-2
; Sequence 2, Application US/11014071
; Publication No. US20050246796A1
; GENERAL INFORMATION:
; APPLICANT: Cigan, Andrew M.
; APPLICANT: Fox, Timothy W.
; APPLICANT: Hershey, Howard P.
; APPLICANT: Unger, Erica
; APPLICANT: Wu, Yongzhong
; TITLE OF INVENTION: Dominant Gene Suppression Transgenes and
; TITLE OF INVENTION: Methods of Using Same
; FILE REFERENCE: 1554
; CURRENT APPLICATION NUMBER: US/11/014,071
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: 60/530,478
; PRIOR FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/591,975
; PRIOR FILING DATE: 2004-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)...(1092)
; OTHER INFORMATION: P95
US-11-014-071-2

Query Match 5.2%; Score 68; DB 11; Length 1092;
Best Local Similarity 57.3%; Pred. No. 0.0063;
Matches 142; Conservative 0; Mismatches 105; Indels 1; Gaps 1;

```
Qy        670 ATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATA 729
              | |||  ||   |||  | | | | | ||  |   | |      || |   ||  ||||
Db        391 ACATAGTTTAGATTCAGTCAGCAAAAAGCTAACATTATGTTTGCTTCTTTTATTCATATA 450

Qy        730 AAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATT 789
              |  ||   ||  ||     |   | |     | |    ||  |  | ||| || | |
Db        451 GTTTTCTTAGCATGAAATTTAAATTCTATATAGTACTGGTTTTTTACGAGCTTACTTAAT 510
```

```
Qy        790 TTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
              ||| | || |||| ||||| |||| ||| || ||  |||||| |||||  ||  ||
Db        511 TTTATTAGGGCTAATTTGGTAACCACATTT-TTCCACGGAATTTCAATTTTCCTAAGGAA 569

Qy        850 AATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACT 909
              |||||||| |||||| ||| |||||||||| | | | |||||  ||| ||| |||
Db        570 AATTAGTTAATTTTCGCTTGGGAAAATAGAAATTTCATGGGAAAATGCGGTTCCCAAACT 629

Qy        910 AGCCCTAG 917
              |||| |||
Db        630 AGCCTTAG 637


RESULT 2
US-10-240-708-36
; Sequence 36, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION:  by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/240,708
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 36
;   LENGTH: 9347
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-36

  Query Match             4.4%;  Score 57.8;  DB 8;  Length 9347;
  Best Local Similarity   48.9%;  Pred. No. 0.64;
  Matches  155;  Conservative    0;  Mismatches  162;  Indels    0;  Gaps    0;

Qy        547 TATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATAT 606
              | |||      || |  ||  || | |   ||| | | || || |||  |  ||| | |
Db       8530 TTTTTATTTTGATAGGAAGGATTTAAATATTTTTATATATTTAGGAGAAATGAATATTTT 8589

Qy        607 CGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTT 666
              |  | |  || ||  ||   || ||| ||| |  || | | | ||||||| |||||
Db       8590 AAGTATTTTAAATAGGATTTTGGTATATTTTTTTTATTTTTTTAGGTTTTTTGTTTTTGTT 8649

Qy        667 TTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACT 726
              | |  ||  | |||  |   ||    || |  |||||| ||    ||| ||  | |||
Db       8650 TATTAATTTAATTTTATAGTTTTTTTTATATAGATTTTGTTTTTGTTTAATTTATTGATA 8709

Qy        727 ATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTG 786
              | | |  || || | |    | | | |      |   | | |          ||| |
Db       8710 TAATATTTTTTGTTGTTTTTATGAATGTAATGTTTTTTTCGTATTGTTTTTTGTTTTGTT 8769

Qy        787 ATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAA 846
              ||| ||   |  |  |  |    | |   ||  |||   ||||  |||||| | |
Db       8770 ATTGTTAGTTAAAGGAAAGTTATTGGTTTTTGGATTGTAAATTTTTTTTTTTTTTTTAGA 8829

Qy        847 AAAAATTAGTTTATTTT 863
              || |||   || | ||
Db       8830 TAAGATTTTATTTTGTT 8846


RESULT 3
US-10-330-773-534
; Sequence 534, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 534
;   LENGTH: 254396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-330-773-534

  Query Match             4.2%;  Score 55.6;  DB 7;  Length 254396;
  Best Local Similarity   46.2%;  Pred. No. 3.3;
  Matches  184;  Conservative    0;  Mismatches  214;  Indels    0;  Gaps    0;

Qy        452 AAAAAAATAACCAAAGTAACTAATCCACTCGAAAGCTATCATGTAATGTTTAAAGAAACA 511
              ||| ||| ||  || | ||  ||    |   ||  |     |  ||     ||  || |
Db     201251 AAAGAAAAAAGAAAGGAAAAGAAAAGAGGAGAGGGAAGAAAGAAAAGAAAAAAGAAAGGA 201310

Qy        512 TCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTT 571
                 || |  | |||| ||| | |||| | ||   |           |  | ||  ||
Db     201311 AAAGAAAAGAAAAAGATCTTCTGACAAAATAGGCCATTCATTGGGCCAAACATTTTAGTG 201370

Qy        572 ACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGA 631
              | | |    || || || |  | | |||||        |    || | || |  |||
Db     201371 GCTGGTCTGGAATATTTATGGACTATAAATTTATCTTTAGTTCAAACCATCAGATTCATG 201430

Qy        632 TTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACA 691
              ||| |  |||| ||  ||   |||| | | | | ||||||   |  |  |    || |
Db     201431 TTTATTCTTTTAATATTTACCATTTATTAACTATATTTATACTAACATAATGACTTTTTA 201490

Qy        692 ATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCA 751
              | | | | |||| ||         |   || |    |||||||  ||||    |||  |
Db     201491 AAACTGGGGTTTTTTTTTTCAGCAACTTAAATAAAGAAAAGTCAGCTTTACTACAGACGA 201550

Qy        752 CGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAA 811
              | ||   | |  |   || | ||   |||| | |  |     |    |||   |
Db     201551 AGAATCAAATGGCATTAGATTTAGAACAACTTCTAAATCATCGTGAGGTATTTTTCCTAT 201610

Qy        812 CCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
                  |  ||  ||  |||||    ||||  |||||
Db     201611 TTTGTCATAATTGTAATCATTTTAGGCTTTTTTAAAAA 201648


RESULT 4
US-09-925-065A-54272
; Sequence 54272, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
```

```
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 54272
;   LENGTH: 915
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-54272

  Query Match             4.2%;  Score 54.8;  DB 6;  Length 915;
  Best Local Similarity   51.0%;  Pred. No. 1.2;
  Matches  151;  Conservative    1;  Mismatches  143;  Indels    1;  Gaps    1;

Qy        596 ACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATT 655
Db        347 AAAGATACTATATATATTTATTATATATATTTTATATATTTATTATATATAATATATATT 406

Qy        656 TTGTTATTGTTTTTATATACATTTTCT-TCTCTTACAATAGAGTGATTTTCTTCCGATTT 714
Db        407 ATATATAATATATATTATATATATTTTATATATTATATATAATATATATTATATATATAATA 466

Qy        715 TATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCA 774
Db        467 TATAATATATATTATATATTTTATATATWATATATAATATATATTATATATTTTATATTA 526

Qy        775 AAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTT 834
Db        527 TATATAATATATATTTTTATATAATATATATTATATATATTTTATATATTATATATAATATA 586

Qy        835 GATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGA 890
Db        587 TATTTTTTATATAATATATATTATATATTTTATATATATTTTATATATTATATATA 642

RESULT 5
US-09-925-065A-803046
; Sequence 803046, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 803046
;   LENGTH: 488
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-803046

  Query Match             4.1%;  Score 53.6;  DB 6;  Length 488;
  Best Local Similarity   49.3%;  Pred. No. 1.7;
  Matches  137;  Conservative    1;  Mismatches  140;  Indels    0;  Gaps    0;

Qy        610 AAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTT 669
Db         37 ATGACATGGTCAGACTTTCTGATTTGTTCTGAACATCCTTTCTTTTTAAACAACCAGTTA 96

Qy        670 ATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATA 729
Db         97 TTTTATTTTAGGACTAAACTTACTACACAGGATTCTTTCTTATATAATAGTATTTCTCTT 156

Qy        730 AAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATT 789
Db        157 TAAACTCTTTTTACCTCAAAAACTACCTCTTTACAACTTTCTTTACATTTTTTATTTC 216

Qy        790 TTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
Db        217 CTGGTTCCTTTTACTTTGTTTTATATATAACTTTAAATAAGCTTTGAATTAGACAAAACA 276

Qy        850 AATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTT 887
Db        277 TGTTCTTTTTTTTTTTTTTTTAGAAAAAAAARACACTT 314

RESULT 6
US-10-240-708-73/c
; Sequence 73, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION:  by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/240,708
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 73
;   LENGTH: 5152
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-73

  Query Match             4.1%;  Score 53.6;  DB 8;  Length 5152;
  Best Local Similarity   48.2%;  Pred. No. 3;
  Matches  219;  Conservative    0;  Mismatches  224;  Indels   11;  Gaps    2;

Qy        425 CTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGAA 484
Db       4137 CTACAATCCCAACTACTCAAAAATCTCAAACAAAAAAATAACTTAAACCCAAAAATTCAA 4078

Qy        485 AGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAG 544
Db       4077 AACTATCATAAACCGTAATTACGCCACTACATTTCAACCTAAACAAAAAACAAAACCCTA 4018

Qy        545 CATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTAT 604
Db       4017 TCTCAAAAAAAAAAAAAAAAAAAAAAAAAAATTTAAAACAATAAATAACAACGCCATCA 3958

Qy        605 ATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTG 664
Db       3957 TATAACTATTATTATACTTATATTAAAAAATACTTTAAAATTTTTTTATTTTTTATTTTT 3898

Qy        665 TTTTTATATACATTTT-CTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATG 723
Db       3897 TATTTTTAAAATTTTAACTTATTTTAAATTCAAAAAATATATATACAAATTTATTACACG 3838

Qy        724 ACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTT 783
```

```
Db       3837 AACATAAAATCAATTTTCTTTAACACC----------TAAAATTAAAAAATAAAAAATAT 3788

Qy        784 CTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTC 843

Db       3787 TTTCATTAAAACTAATTTCCTTTAACACAATCTTTAATTTATAAAAATTTTCCCATTTTT 3728

Qy        844 AAAAAAAATTAGTTTATTTTCTCTTTATAAAATA 877

Db       3727 AAAAAAAATTCCTAAAATACCTAAATATTAAAAA 3694


RESULT 7
US-10-240-708-63
; Sequence 63, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION:  by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/240,708
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 63
;   LENGTH: 5562
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-63

  Query Match             4.1%;  Score 53.4;  DB 8;  Length 5562;
  Best Local Similarity   52.4%;  Pred. No. 3.3;
  Matches  140;  Conservative    0;  Mismatches  126;  Indels    1;  Gaps    1;

Qy        614 TAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTTATAT 673

Db       1325 TTATATTTTAAGGTTATTATTTATTTTTTGTTTGGGGTGATTTTTTTTTTTGTTTTTTTTA 1384

Qy        674 ACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGT 733

Db       1385 TTATTTTAATATTTTATTATTATTTTTTTTTTTTCGGTTAGGAATGATTTTTTTTTGAT 1444

Qy        734 CATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATC-TTTCTGATTTTT 792

Db       1445 TATATTTATATTTTTTTATTTTAGTTTAGTTTTTTTTTTTTTTTATTTATTTTTTATTATT 1504

Qy        793 TTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAAT 852

Db       1505 TTTATTTTTTGTTAGTTAGATTTTTTTTTTTTTATTATTTTTTTGTTTTTTATAGTGTTT 1564

Qy        853 TAGTTTATTTTCTCTTTATAAAATAGA 879

Db       1565 TTTTTTATTTTTTTTTTGTTTAAGAAA 1591


RESULT 8
US-10-995-561-13314/c
; Sequence 13314, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13314
;   LENGTH: 119036
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(119036)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-2
US-10-995-561-13314

  Query Match             4.1%;  Score 53.2;  DB 8;  Length 119036;
  Best Local Similarity   48.4%;  Pred. No. 7.4;
  Matches  177;  Conservative    0;  Mismatches  188;  Indels    1;  Gaps    1;

Qy        619 TATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATT 678

Db      31623 TAAAATCTTATTTTATATATATTTTATATATTTTATATATAAAATCTTATTTTATATATA 31564

Qy        679 TTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTT 738

Db      31563 TTTTTATATATAAAAAATCATATTTTATATATATTTTATATATTTTATATATATGAAATC 31504

Qy        739 TTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGA 798

Db      31503 TTATTTTATATATTATTTTATATATATTTTATATATATATAAAATCTTATTTTATATATA 31444

Qy        799 GCTAGTTTGGCAACCCTGT-TTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTT 857

Db      31443 TTTTATATATATATAAAATCTTATTTTATATATATTTTATATATATATAAAATCTTATTT 31384

Qy        858 TATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAG 917

Db      31383 TATATATATTTTATATATATATAAAATCTTATTTTATATATATTTTATATATATATAAAA 31324

Qy        918 AATGTTTTCCCAATAAATTACAATCACTGTGTATAATTATTTGGCCAGCCCCATAAATTA 977

Db      31323 TCTTATTTTATATATATTTTATATATATATAAAATCTTATTTATACAGCCATTTTTCTAA 31264

Qy        978 TTTAAA 983

Db      31263 ATAAAA 31258


RESULT 9
US-09-925-065A-803045
; Sequence 803045, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
```

```
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 803045
;   LENGTH: 488
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-803045

  Query Match           4.0%;  Score 52;  DB 6;  Length 488;
  Best Local Similarity   48.9%;  Pred. No. 3.3;
  Matches  136;  Conservative    1;  Mismatches  141;  Indels    0;  Gaps    0;

Qy        610 AAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTT 669
              | |  | | |  ||| ||| ||||| |  |   ||| |||  | |||     |     ||
Db         37 ATGACATGGTCAGACTTTCTGATTTGTTCTGAACATCCTTTCTTTTTAAACAACCAGTTA 96

Qy        670 ATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATA 729
              |.||  ||          | || |  |  ||  |||  || | |||  ||  || | |
Db         97 TTTTATTTTAGGACTAAACTTACTACACAGGATTCTTTCTTATATAATAGTATTTCTCTT 156

Qy        730 AAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATT 789
              :  |  |||| ||     |  |  || |     ||   |  | | |||     |
Db        157 TMAACTCTTTTTTACCTCAAAAACTACCTCTTTACAACTTTCTTTACATTTTTTTATTTC 216

Qy        790 TTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAA 849
              |  |      || ||||       | |  |||| ||  |  ||||| ||   |||||
Db        217 CTGGTTCCTTTTACTTTGTTTTATATATAACTTTAAATAAGCTTTGAATTAGACAAAACA 276

Qy        850 AATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTT 887
                ||  ||| |||| | ||| || || | || ||||||
Db        277 TGTTCTTTTTTTTTTTTTTTTAGAAAAAAAGACACTT 314


RESULT 10
US-10-240-708-11/c
; Sequence 11, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION:  by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/240,708
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
;   LENGTH: 6317
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-11

  Query Match           3.9%;  Score 51.6;  DB 8;  Length 6317;
  Best Local Similarity   45.8%;  Pred. No. 7.1;
  Matches  218;  Conservative    0;  Mismatches  254;  Indels    4;  Gaps    1;

Qy        424 ACTATTGGCACTACCCAGAACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGA 483
              || ||    |  |   | || || |  |||||||    | |||  |    ||  | | |
Db       2030 ACGATAATAAAAAAAAAAAAAAAAAAAAATAAAAAAAACTCAAAAACACTATATAAAAACCA 1971

Qy        484 AAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAA 543
              ||   |  || |||  |||||  |||      ||| ||| ||  ||      |||||||||
Db       1970 AATACAATATATAAACTTTAATTAAAATCTAATTCAAAACAAAATAACTATAAAAAACAT 1911

Qy        544 GCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTA 603
              |    |  | | | | |||  | | |    |  | || ||  || |    | ||| |
Db       1910 TAAACAATAAATAAAAATAAAAAAATTAATTTTCAAATAATATTAAAAAATTATAAAATT 1851

Qy        604 TATCGAAAGGTAAGCT----ATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGT 659
              |    |||  | |         || |  |  | || ||  |  | | ||||   || |
Db       1850 TTCTTAAATATTATAATATAATTATATCTATATAATTTATCCTTACTCTTAAAATATACA 1791

Qy        660 TATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAA 719
              |||      |  | | | |   | ||||| ||||     |  | |||    | ||  |||
Db       1790 TATCAAAATATTTAAAAATAAAATATCTTATAATATTTATAAAATACTTTAAAATTATTAA 1731

Qy        720 AATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAAT 779
              |  | ||    | |  |  |  | |  |      ||     |  |  |  || ||
Db       1730 TCATAAAAAAATTAAAATAATTAAATACATACGACAAAATATTAACAATTATTAAATCTA 1671

Qy        780 CTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTT 839
              ||  | | | || |     | |         | |  | | ||| |  |    |  |  ||
Db       1670 ATTCATTAATATATTTATTAACAATACTCTTTCAATTTCTCTATATATTTAAAATATTTC 1611

Qy        840 TTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAAT 895
              |  ||||| ||| |    | ||   ||| |||||||     | | |   |||| |
Db       1610 ATAAAAAAATAATAAAAAAACTTAAACTTAATAAAATTACTTAAATTCTAAAAACT 1555


RESULT 11
US-10-330-773-242/c
; Sequence 242, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 242
;   LENGTH: 355211
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(355211)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-242

  Query Match           3.9%;  Score 51.6;  DB 7;  Length 355211;
  Best Local Similarity   46.1%;  Pred. No. 18;
  Matches  211;  Conservative    0;  Mismatches  244;  Indels    3;  Gaps    1;

Qy        443 ACAAGATTTAAAAAAATAACCAAAGTAACTAATCCACTCGAAAGCTATCATGTAATGTTT 502
              |||  || |||| ||   ||  ||   |   | | | | | |  |     |
Db      61794 ACACAATATAAATAATATACATAACATATATGTATATTTTATATATGTATATTTTATATA 61735

Qy        503 AAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTTCGAAAGAGACA 562
              |    | || | ||| |   | |    | ||| | |   |  |||| | | |    | |
Db      61734 TATTTTATATATTTTATATATATGTATATTTTATATATAATATATATATTGTATATAATA 61675

Qy        563 AATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAAGGTAAGCTATG 622
              |  |||  || | | ||    || ||  |   | | | |||||  | |  ||      |
Db      61674 ATATATAATATATTATA---TTATATATAATATATATAATATATATATAAATATATATTA 61618

Qy        623 ACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTATATACATTTTCT 682
              | | || | |  | | | | || |   | |  |||   |  | ||||||  || | |
Db      61617 TATATAATATGTATAATATATAATATTTTATATATAATATGTATAATATATATTTTATAT 61558

Qy        683 TCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAAATGACTATAAAGTCATTTTTAT 742
```

```
Db      61557 ATAATAATATGTACAATATATATTTTATATATAATAATATGTACAATATATATTTTATAT 61498

Qy        743 ATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTA 802

Db      61497 ATAATAATATGTACAATATATATTTTATGTATAATATGTATAATATATATTTTATGTATA 61438

Qy        803 GTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAAAAATTAGTTTATTT 862

Db      61437 ATATATATTTTATGTATAATATATATTTTACGTATATTTTATATATAATATATAATATTT 61378

Qy        863 TCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGT 900

Db      61377 TATATATAATATATAACATTTTATATATAATATATAAT 61340


RESULT 12
US-11-121-086-25
; Sequence 25, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
;  APPLICANT: POULSEN, TIM S.
;  APPLICANT:  NIELSEN, KIRSTEN V.
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;  FILE REFERENCE: 09138.6000-00000
;  CURRENT APPLICATION NUMBER: US/11/121,086
;  CURRENT FILING DATE:  2005-05-04
;  PRIOR APPLICATION NUMBER: 60/567,570
;  PRIOR FILING DATE: 2004-05-04
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
;   LENGTH: 173602
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-25

  Query Match             3.9%;  Score 51.4;  DB 12;  Length 173602;
  Best Local Similarity   45.5%;  Pred. No. 17;
  Matches  219;  Conservative    0;  Mismatches  261;  Indels    1;  Gaps    1;

Qy        474 ATCCACTCGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCT 533

Db     136539 ATATATTATATATTTATATGTTTATGTATATATAAATATATAAT-ATATAAATATATTTA 136597

Qy        534 TAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAG 593

Db     136598 TATATTATATATTTATATAATATTTTATATAATATGTTATATATTTTTATATAATATATT 136657

Qy        594 CGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTA 653

Db     136658 ATATATATTTATATGCTATATTATATAGTATATTATATATAATATATTTTATATATTATA 136717

Qy        654 TTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATT 713

Db     136718 TATATATTTATATAATATTTATATATTATATTTATATAATATGTTTATATATTATTATAT 136777

Qy        714 TTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTC 773

Db     136778 TTATATAATATGTTTATATATTATTATATTTATATAATATGTTTATATATTATTATATTT 136837

Qy        774 AAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTT 833

Db     136838 ATATAATATGTTTATAATATATATAATATATTTATATAATATGTTTATATATTATATATA 136897

Qy        834 TGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAA 893

Db     136898 TAATATATTTATATAATATGTTTATATATTATATATATAATATATTTATATTATTATA 136957

Qy        894 ATAGAGTTGCCAGACTAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAA 953

Db     136958 TATATTTTATAATATATAATATATATAATATAATATATATATTATATATATTATATATAT 137017

Qy        954 T 954

Db     137018 T 137018


RESULT 13
US-10-330-773-26
; Sequence 26, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
;   LENGTH: 687411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(687411)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-26

  Query Match             3.9%;  Score 51.2;  DB 7;  Length 687411;
  Best Local Similarity   46.2%;  Pred. No. 25;
  Matches  170;  Conservative    0;  Mismatches  198;  Indels    0;  Gaps    0;

Qy        531 TCTTAAAAAACAAGCATATTTCGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAA 590

Db         98 TATATAATAATTAAATTATAGTAAAAATATTTATTTATAATATAAATATAAAATGTATAT 157

Qy        591 GAGCGACAAATTATATCGAAAGGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTG 650

Db        158 TATATTATAATTATATATAATTATATTGTAATTAAATTATATATAATTTTATATATTATA 217

Qy        651 TTATTTTGTTATTGTTTTTATATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCG 710

Db        218 TATTATATATATTAAATTATAATACAGCATATTGTAATACAATATATTTTATATTAATAT 277

Qy        711 ATTTTATAAAATGACTATAAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCG 770

Db        278 AGTATAGTAAAATATATAATATATAATATAATTTCATTATAATAAAATATAATAAAAATA 337

Qy        771 TTCAAAAATCTTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAA 830

Db        338 TTTAAAAATGTATTTTATATCTTACATATATATTCAAATATATACATATATTTATATAAA 397

Qy        831 TTTTGATTTTTTCAAAAAAAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGA 890

Db        398 TATATACTTAATATTTATATATTAATATATATTTTTATAAATATATATATATTATATAAA 457

Qy        891 AAAATAGA 898

Db        458 TATAAATA 465


RESULT 14
US-09-925-065A-667262
; Sequence 667262, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
```

; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 667262
; LENGTH: 677
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-667262

Query Match 3.9%; Score 51; DB 6; Length 677;
Best Local Similarity 45.6%; Pred. No. 5.4;
Matches 214; Conservative 1; Mismatches 251; Indels 3; Gaps 1;

```
Qy        492 ATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTT 551
              || |:|| | |||| |     ||| ||   | |   | ||| | |   |   ||||
Db         31 ATATWATATATAAAATATATAATATATAATATATGTATATTTTATATATAATATATATAA 90

Qy        552 CGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAA 611
                | | |    |  ||   || |  ||  |   || ||  | | | ||| ||| | |
Db         91 ATATATATTATATATATAATATATATTTTATTTATATATAATATATATTTTATTTATA 150

Qy        612 GGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTAT 671
                |||  |||    || | | | | |  ||  || |  |||| | |   | | ||
Db        151 TATAAAATATATATTTTATTTATATATAATATATTTATATTTTTTATTTATATATAAT 210

Qy        672 ATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA---ATGACTAT 728
              ||| ||||   | |  || | || |   ||||   |  || |  || |   ||   |||
Db        211 ATATATTTATATATTATATATTATATATATTTATATTATATATTATATATATTTATAT 270

Qy        729 AAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGAT 788
              |   |   | ||||||     | | ||    ||  | |  | | |    | |
Db        271 ATTATATATAATATATATATTTATGTATTATATATTATATATATTTATATATGTAATATA 330

Qy        789 TTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAA 848
              | |||    |    | |    | | | | | | | |  ||  ||   | ||
Db        331 TATTTATGTATTATATATTTATATATTATATATATTATATATATTATATACTTAATATAA 390

Qy        849 AAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGAC 908
              |  || | ||| || | | |||   ||| | |   |  ||   ||| | ||   | |
Db        391 TATATAATATATATTATATATATTGTATATATATAATATATATTTATATATTTTATATAT 450

Qy        909 TAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTAT 957
              |    || | |  | |    |   | || |   |     ||||  |||
Db        451 AATATATATATTTATATATAATATATATAAATATATATATTATATATAT 499
```

RESULT 15
US-09-925-065A-667263
; Sequence 667263, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 667263
; LENGTH: 677
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-667263

Query Match 3.9%; Score 51; DB 6; Length 677;
Best Local Similarity 45.6%; Pred. No. 5.4;
Matches 214; Conservative 1; Mismatches 251; Indels 3; Gaps 1;

```
Qy        492 ATGTAATGTTTAAAGAAACATCTATTAAAACCACGATCCTCTTAAAAAACAAGCATATTT 551
              || | || | ||||  |   : |||  ||   | |   | ||| | |  |  ||||
Db         31 ATATTATATATAAAATATATWATATATAATATATGTATATTTTATATATAATATATATAA 90

Qy        552 CGAAAGAGACAAATTATGTTACAGTTTACAAACATCTAAGAGCGACAAATTATATCGAAA 611
                | | |    |  ||   || |  ||  |   || ||  | | | ||| ||| | |
Db         91 ATATATATTATATATATAATATATATTTTATTTATATATAATATATATTTTATTTATA 150

Qy        612 GGTAAGCTATGACGTTCAGATTTTTCTTTTTCATTCTTGTTATTTTGTTATTGTTTTTAT 671
              |||  |||    || | | | | |  ||  || |  |||| | |   | | ||
Db        151 TATAAAATATATATTTTATTTATATATAATATATATTTATATTTTTTATTTATATATAAT 210

Qy        672 ATACATTTTCTTCTCTTACAATAGAGTGATTTTCTTCCGATTTTATAAA---ATGACTAT 728
              ||| ||||   | |  || | || |   ||||   |  || |  || |   ||   |||
Db        211 ATATATTTATATATTATATATTATATATATTTATATTATATATTATATATATTTATAT 270

Qy        729 AAAGTCATTTTTATATAAGAGCACGCATGTCGTAGATTCTCGTTCAAAAATCTTTCTGAT 788
              |   |   | ||||||     | | ||    ||  | |  | | |    | |
Db        271 ATTATATATAATATATATATTTATGTATTATATATTATATATATTTATATATGTAATATA 330

Qy        789 TTTTTTAAGAGCTAGTTTGGCAACCCTGTTTCTTTCAAAGAATTTTGATTTTTTCAAAAA 848
              | |||    |    | |    | | | | | | | |  ||  ||   | ||
Db        331 TATTTATGTATTATATATTTATATATTATATATATTATATATATTATATACTTAATATAA 390

Qy        849 AAATTAGTTTATTTTCTCTTTATAAAATAGAAAACACTTAGAAAAATAGAGTTGCCAGAC 908
              |  || | ||| || | | |||   ||| | |   |  ||   ||| | ||   | |
Db        391 TATATAATATATATTATATATATTGTATATATATAATATATATTTATATATTTTATATAT 450

Qy        909 TAGCCCTAGAATGTTTTCCCAATAAATTACAATCACTGTGTATAATTAT 957
              |    || | |  | |    |   | || |   |     ||||  |||
Db        451 AATATATATATTTATATATAATATATATAAATATATATATTATATATAT 499
```

Search completed: March 6, 2006, 10:32:43
Job time : 1372.54 secs